Combining Virtual Learning Environments and Intelligent Pedagogical Agents for a New Approach to Classroom Lessons

Sapienza Università di Roma
Dipartimento di Ingegneria Informatica
Automatica e Gestionale Antonio Ruberti

Dottorato di Ricerca in Ingegneria Informatica
XXIX Ciclo

Candidate: Annalisa Terracina
ID number: 673854

Thesis Advisor
Prof. Massimo Mecella

Co-Advisor
Prof. Luigia Carlucci

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Engineering in Computer Science
March 2017
Thesis not yet defended

Combining Virtual Learning Environments and Intelligent Pedagogical Agents for a New Approach to Classroom Lessons
Ph.D. thesis. SAPIENZA Università di Roma

© 2017 Annalisa Terracina. All rights reserved

This thesis has been typeset by LATEX and the Sapthesis class.

Version: June 8, 2017
Website: http://www.dis.uniroma1.it/~aterracina
Author’s email: terracina@dis.uniroma1.it
To my beloved grand mother Carmela Senise.

"Ignorance more frequently begets confidence than does knowledge: it is those who know little, and not those who know much, who so positively assert that this or that problem will never be solved by science." C. Darwin
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Glossary</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>Extended abstract</strong></td>
<td>vii</td>
</tr>
<tr>
<td>Research Context and Background</td>
<td>vi</td>
</tr>
<tr>
<td>Research Objectives and Contributions</td>
<td>ix</td>
</tr>
<tr>
<td>Summary of Publications</td>
<td>xii</td>
</tr>
<tr>
<td>Thesis Structure and Outline</td>
<td>xiv</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 21st Century Skills</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Artificial Intelligence grand challenges in education</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Game Based Learning</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Combining technical and pedagogical challenges</td>
<td>7</td>
</tr>
<tr>
<td><strong>2 Virtual Learning Environment and Intelligent Tutoring System</strong></td>
<td>9</td>
</tr>
<tr>
<td>2.1 Virtual Learning Environments</td>
<td>9</td>
</tr>
<tr>
<td>2.1.1 3D simulations and microworlds</td>
<td>11</td>
</tr>
<tr>
<td>2.1.2 3D environments as interfaces to learning resources</td>
<td>12</td>
</tr>
<tr>
<td>2.1.3 3D multi-user VLEs</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Intelligent Tutoring Systems</td>
<td>14</td>
</tr>
<tr>
<td>2.2.1 Conversational Agent</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2 Intelligent Pedagogical Agent</td>
<td>18</td>
</tr>
<tr>
<td><strong>3 NewEarth, a novel approach to Game Based Learning</strong></td>
<td>23</td>
</tr>
<tr>
<td>3.1 Pedagogical aims</td>
<td>23</td>
</tr>
<tr>
<td>3.2 Serious Game Design</td>
<td>31</td>
</tr>
<tr>
<td>3.3 Research open questions</td>
<td>35</td>
</tr>
<tr>
<td><strong>4 NewEarth architecture &amp; design</strong></td>
<td>37</td>
</tr>
<tr>
<td>4.1 VLE</td>
<td>40</td>
</tr>
<tr>
<td>4.1.1 Presentation tier: teachers’ workspace</td>
<td>40</td>
</tr>
<tr>
<td>4.1.2 Presentation tier: Students’ workspace</td>
<td>43</td>
</tr>
<tr>
<td>4.1.3 Logic Tier</td>
<td>44</td>
</tr>
<tr>
<td>4.1.4 Data Tier</td>
<td>50</td>
</tr>
<tr>
<td>4.2 IPA</td>
<td>52</td>
</tr>
<tr>
<td>4.2.1 Dialogue Management System</td>
<td>53</td>
</tr>
<tr>
<td>4.2.2 Emotions and schooling</td>
<td>55</td>
</tr>
</tbody>
</table>
4.2.3 Unsolicited Hints .............................................. 59

5 Realizing NewEarth ............................................. 65
  5.1 VLE ................................................................. 65
    5.1.1 Presentation tier: Teachers’ workspace ................. 65
    5.1.2 Presentation tier: Students’ workspace ................. 71
    5.1.3 Logic Tier ................................................... 76
    5.1.4 Data Tier ................................................... 81
  5.2 IPA ................................................................. 82

6 Game evaluation & learner assessment ......................... 86
  6.1 Introduction .................................................. 86
  6.2 Ecological test run no. 1: Emotional Intelligent Virtual Tutor in a stand-alone app ............................................................. 87
    6.2.1 EIVT and Android app ..................................... 87
    6.2.2 Ecological Test ............................................. 89
    6.2.3 Discussion of Results ..................................... 93
  6.3 Ecological test run no. 2: First version of NE ................ 94
    6.3.1 Ecological environment description .................... 94
    6.3.2 Learning assessment ..................................... 100
  6.4 Ecological test run no. 3: Improved version of NE ........... 108
    6.4.1 Ecological environment description .................... 108
    6.4.2 Learning Assessment ..................................... 112
  6.5 Unsolicited hints .............................................. 120

7 Conclusions ....................................................... 123
  7.1 New vision in education ..................................... 123
  7.2 Lessons learned ............................................. 125
  7.3 Future works ................................................ 126

Appendices .......................................................... 129

A Emotion & schooling project plan .................................. 130

B Learning contents ................................................. 135

C Additional work ................................................... 137
Glossary

3DVG 3D Virtual Game.

AI  Artificial Intelligence.

API  Application Program Interface.

CTM  Cognitive Trait Model.

DMS  Dialog Management System.

HCI  Human Computer Interaction.

IAP  Interrogative/Adjective Pronoun.

ICT  Information and Communication Technology.

IDF  Inverse Document Frequency.

IPA  Intelligent Pedagogical Agent.

ITS  Intelligent Tutoring System.

IVT  Intelligent Virtual Tutor.

KDB  Knowledge Database.

KST  Knowledge Space Theory.

LeS  Learning Situation.

MMORPG  Massive Multiple Online Role Playing Game.

NLP  Natural language Processing.

NPC  Non Player Character.

PA  Professor App.

PVB  Professor Virtual Board.
Glossary

REST  Representational State Transfer.
RPG  Role Playing Game.
SG  Serious Game.
STEM  Science, Technology, Engineering, Math.
TEL  Technology Enhanced Learning.
TF  Term Frequency.
VC  Virtual Character.
VE  Virtual Environment.
VLE  Virtual Learning Environment.
VT  Virtual Tutor.
Extended abstract

Research context and background

Europe has experienced a dramatic decline in the number of young people taking Science, Technology, Engineering and Mathematics (STEM) subjects in school and the uptake of careers in the same fields. Thus, it is necessary to rethink how education can be reshaped so that it reconsiders the requirements for new competencies. In this way, it is able to be tailored for “generation.com” or the new “millenials”. A learners’ centred pedagogy using ICT and interdisciplinary research makes personalised and adaptive learning possible. The development and integration of robust and fit-for-purpose digital technologies for learning are crucial to boost innovation in Educational Technologies.

The recent interest in applied sciences such as Information and Communications Technology (ICT) and Biotechnology has not been enough to offset the STEM’s fall, especially among European women. With the current European STEM population ageing, a lack of talent to replace them could have wide-ranging and serious consequences. A shortage of young people going to university to read these subjects, adversely affects the STEM industry and employment markets. If the level of those qualified and literate in STEM subjects drops to a critical level, the whole future of the European knowledge-based society and economy could be at stake. To contrast this tendency there are many initiatives. Among them, the European Commission indications\(^1\) are:

“\textit{A learners’-centred pedagogy using ICT and interdisciplinary research makes personalised and adaptive learning possible. For instance, learning analytic tools support teachers’ work by providing individual feedback and recommendations to students and serious educational games are powerful tools to engage kids (and students of all ages) in learning.}”

Games are structured contexts, with clearly-defined rules, where players must overcome challenges and face opponents (real or game characters) to achieve victory. Games can offer incredibly immersive and engaging environments where users learn

by doing and improve skills and competences related to decision making, strategy, teamwork, social skills, leadership and collaboration. Serious Games (SG) focus on the design, development, use and application of games for purposes other than entertainment. Education and training are areas that afford the use of Serious Games along with other areas such as health, research, emergency planning, advertisement, military purposes, etc. The most striking effect in the use of Serious Games is an increased user motivation and engagement towards the serious objectives. Interaction and simulation are fundamental tools for this motivation and engagement: providing intuitive and innovative forms of interaction with the game, captures the user interest and, creating intelligent game play, maintains that interest.

“By 2030, SG are anticipated to become a useful and reliable tool for learning and training. Educators and trainers will, according to user-specific needs, be able to select the most suitable games for achieving precise learning objectives by utilizing catalogues and databases of SGs. Authoring tools will promote the customization of content and learning procedures, and will be widely used. SGs will be used to motivate students to explore new topics as well solving exercises efficiently, practice the application of learned theory, and to verify the acquisition of knowledge and skill”.

These are the outcomes of the GALA\(^2\) research project (2014) claimed in the result roadmap.

Research has highlighted motivational factors from computer/virtual games, visualization, and Human-Computer-Interaction (HCI) aspects that can greatly motivate the next-generation learner. With respect to the past, the importance and meaning of pedagogical foundation in educational design is clearer to researchers; the infrastructure of schools has developed a lot during the last decade and finally, we are moving toward a new generation of educational use of games. In spite of the existing evidence of success, there is still a limited use of Serious Games. This has mainly to do with social concerns and stigma about the use of games and serious purposes. Other issues relate to physical and cost barriers, hardware and license cost, access (for online games), maintenance and support. As a consequence, the quality of SGs has not met the expectations of educators and the use of games has not become as general as expected, Kiili & Sylvester in [55]. The limited use is also related to the lack of extended evidence of effective application.

The overall aim of our research is to give a contribution in the area of Technology Enhanced Learning (TEL). In this research we presents a SG that is developed as an immersive 3D Virtual Learning Environment (VLE); VLEs integrate several educational resources including multimedia learning material, communication tools, and recommender systems, among others. 3D VLE adds 3D visualization and the ability of the learner to navigate the 3D virtual environment. The usage of VLE assumes that the learner uses computerized support instead of a human instructor, and thus Virtual Characters (VC) in the form of Intelligent Tutoring System (ITS), have been developed to improve interactivity, trying to compensate for the lack of the human aspect. In addition, VLE can support high degrees of individualization that are not possible in classical classrooms environments.

VLE and ITS are fully described in Chapter 2.

\(^2\)http://www.seriousgamesinstitute.co.uk/applied-research/GALA.aspx
Gea 2: A New Earth

“What is merely play is not serious but play itself contains its own, even sacred, seriousness. Hans-Georg Gadamer [37]”

NewEarth is the brand new SG that we have developed. The game aims to teach Science, Technology, Engineering and Mathematics (STEM) to students that approach for the first time specific topics and thus are new to Physics, Chemistry, etc. Unfortunately, there is not a common approach in teaching, Langen & Dekkers in [58], and even students’ age varies depending on countries. However, we developed the game to be aimed at 13 to 18-year-old students, which is the most appropriate age range in almost all European countries. The role-playing game has been thought of as a supporting teaching tool to be used in classes, as an alternative to the more classical frontal lessons, Michael & Chen in [70].

A NEO (Near Earth Object) will impact the Earth causing earthquakes and tsunamis that will devastate continents, and there is no way to deviate or destroy the asteroid. The team’s challenge is to choose an exo-planet (a planet with characteristics similar to our planet Earth) to colonize among three available ones. The individual challenge depends on the chosen role: the game is a role-playing game (RPG) where each student becomes a player with her abilities and her tasks. Roles have been designed having in mind H. Gardner’s theory of multiple intelligences and entry points [38], see Chapter 4. Each player should solve a task related to the habitability of the planets. A screen-shot of the SG from student’s perspective is depicted in Figure 0.1(a) while another screen-shot from teacher’s perspective is depicted in Figure 0.1(b).

A full description of the SG, from users’ perspective, can be found in Chapter 3, Section 3.2.

Research objectives and contributions

As already discussed, Europe has experienced a dramatic decline in the number of young people taking STEM subjects in school. A learners’ centred pedagogy using ICT and interdisciplinary research makes personalised and adaptive learning possible. The development and integration of robust and fit-for-purpose digital technologies for learning are crucial to boost innovation in Educational Technologies. That leads us to the first research objective:
Objective 1.
Provide a contribution in the area of Technology Enhanced Learning maturing STEM acquisition skills.

While there is an increasing interest in Serious Games, Virtual Learning Environments, Intelligent Tutoring Systems, etc., there has not been yet an evidence in combining all these aspect into one. That is probably due to the fact that each one of the afore mentioned single research is quite complex to implement. Our intuition is to lower the complexity of the single system and to compose them together to have a global system that can easily manage different aspects of Technology Enhanced Learning. That idea leads us to introduce the second research objective.

Objective 2.
Design and implement a Serious Game that makes use of Virtual Learning Environment and Intelligent Tutoring system in the form of Intelligent Pedagogical Agent (IPA).

When the virtual environment becomes complex with many places to visit and learn from, pedagogical scenarios are pervasive and guidance is needed by an agent to aid the learner to find help relevant to the educational goals. IPA, can provide intelligent guidance for the learner in the immersive VLE. The use of avatars to only provide a visual representation of players character, although great for improving motivation support, cannot be enough because avatar actions are guided by the player himself/herself. Instead a pedagogical pro-active and autonomous agent is needed to assist or do other pedagogical actions with or on behalf of the learner in the environment. It requires further intelligence abilities that can improve pedagogical functions and act intelligently on behalf or with the learner to achieve goals.

AI affects growth and productivity in many sectors (for example, transportation, communication, commerce, and finance). However, one exception is education; specifically, very few AI-based learning systems are consistently used in classrooms or homes. Yet the potential exists for AI to have a large impact on education, as stated by Woolf et al. in [110]. Classrooms and textbooks are especially inappropriate for people who use mobile and digital technology every day. For example, digital natives learn and work at twice speed, through parallel processing, and connected to others as explained by Beavis in [5]. For digital natives, information is instantly available, change is constant, distance and time do not matter, and multimedia is omnipresent.

To mentor effectively and support individuals or groups while learning, an intelligent system needs to model the changes that occur in learners. Estimates of a learner’s competence or emotional state, stored in user models, represent what learners know, feel, and can do. Mentoring systems should also support

---

3 The term digital native was coined and popularized by education consultant Marc Prensky in his 2001 article entitled Digital Natives, Digital Immigrants, in which he relates the contemporary decline in American education to educators’ failure to understand the needs of modern students.
learners with their decision making and reasoning, especially in volatile and rapidly changing environments. Learners need to make informed decisions and justify them with evidence, gathered through collaboration and communication. ITSs with conversational dialogue form a special category of educational technologies. These conversational ITSs are based on explanation-based constructivist theories of learning and the collaborative constructive activities that occur during human tutoring, see Rus et al. in [89].

**Objective 3.**

Design and implementation of an Intelligent Pedagogical Agent capable of conversation in Natural Language and of providing unsolicited hints during the game play, exploiting ad-hoc developed AI algorithms.

In the context of autonomous intelligent and adaptive systems, classifying and modelling the learner plays a crucial role and bears one of the most substantial challenges to research. In fact, classifying goes back to the first attempts of assigning grades to students. In General, classifying learners was the subject of a large body of research (see Sampson & Karagiannidis in [91]) and has a very long tradition. The cognitive aspects, are in the centre of most attempts and frameworks, for example in the CTM, the Cognitive Trait Model, which enables student modelling on the basis of cognitive abilities and resources, as explained by Lin & Patel in [64]. A well-elaborated and validate framework for assessing in particular cognitive aspects comes from Kickmeier-Rust & Albert [54], utilizing the ideas of CbKST which is an extension of the originally behavioural Knowledge Space Theory KST by Doignon & Falmagne in [30] where a knowledge domain is characterised by a set of problems or test items/tasks. The knowledge state of a learner is identified on the subset of problems s/he is capable of solving. Due to mutual dependencies between the items captured by prerequisite relations, not all potential knowledge states are supposed to occur. The set of all possible states is called a knowledge structure. To account for the fact that a problem might have several prerequisites (i.e., and/or-type relations) the notion of a prerequisite function was introduced. That leads to the fourth objective of our research.

**Objective 4.**

Game evaluation and learners assessment involving stakeholders, namely students and teachers, during formal teaching activities.

The principal idea of CbKST is to separate the observable behavioural aspects, i.e., whether a learner masters a problem or test item, from the not directly observable construct of aptitude/ability/knowledge behind the performance. The entities of aptitude matching the concept of problems or items are called skills or competencies. Equal to knowledge structures, prerequisites between competencies establish competence structures that include only meaningful sets of competencies a person can have. The relationships between the competencies and problems/items are established by a skill function. Such function assigns a collection of subsets of
competencies (i.e., competence states) to each problem that are relevant for solving it. By associating competencies with the problems/items of a domain, a knowledge structure on the set of problems is induced. The latent competencies can be uncovered on the basis of a person’s observable performance. To account for the specific needs of educational games and to achieve a non-invasive, unobtrusive assessment, Kickmeier-Rust & Albert in [54], developed a formal model of the problem solving behaviour in game-based Learning Situations (LeS). Basically, LeS are characterised by a large degree of freedom and complex problem solving demands. The problem solution process is considered to be a meaningful sequence of problem solution states establishing a problem space. Stochastic process models are applied in order to estimate the likelihood of certain state transitions and to estimate the probability of reaching a solution state (within a specific time interval). In other terms, a LeS is segmented in to a set of possible problem solution states, each mapped to one of a set of possible competence states. In this way, the educational AI of a game can interpret the behaviour of the learner in terms of available knowledge, un-activated knowledge, or missing knowledge, simply by mapping the actions of the learner to competence states. Practically speaking, the global framework, termed "micro adaptively", attempts to analyse a learner’s progress in the game environment and to associate a probability of available or lacking competencies on a probabilistic level. Depending on an increase (what actually is desired) or a decrease of the probability of specific competencies, pedagogical/didactic meta-rules are utilised to select a specific interventions and feedback.

Summary of publications

The main contributions and results discussed in this thesis have been presented in the following published works. Figure 0.1 gives an overview of the publications in relation to the four objectives discussed in the previous section.


**Figure 0.1.** Papers and research’s objectives.


**Dissemination**

The research outcomes have been disseminated in specific events devoted to learning, science and school.
[E1] *Science On Stage festival*, 2015. The research presented was ranked in the first 10 *learning innovations* and has been selected, by a national committee, to be one of the ten Italian representative at the European festival of innovating science.

[E2] *European Geoscience Union (EGU) - The Solar System and Beyond*, 2016. The research presented was ranked in the 5 first *learning innovations* and has been selected, by a national committee, to be one of the five Italian representative at the teachers workshop held inside the EGU general assembly.

In the following section we give an overview of the thesis structure and outline.

**Thesis structure and outline**

This thesis is structured as follows.

- **Chapter 1** provides an overview and introduction to the broad field of Enhanced Technology Learning or, generally speaking, Educational Software. The importance of the 21st Century skills that underpin this field of research, defined by UNESCO’s Task Force on education, are introduced as well. We present in this chapter also the Artificial Intelligence grand challenges in education as well as an introduction to Serious Games. In fact, the aim of the chapter is to introduce the reader to the challenges that guided the research presented in this thesis.

- **Chapter 2** elaborates on the challenges related to Virtual Learning Environment and Intelligent Tutoring System. First, we give a definition of these systems and then we present an overview of the most research’s achievements in this field.

- **Chapter 3** gives an overview on the pedagogical reasoning that leads us to some choices during the design of the SG. It gives also an introduction to some principle of game design. We also introduce some research questions still open that we have addressed in this research study.

- **Chapter 4** introduces the design of *NewEarth* SG; the SG’s design has been done trying to tackle technical and pedagogical challenges as discussed in Chapter 1.

- **Chapter 5** focuses on the implementation of *NewEarth* SG; a detailed explanation of the SG architecture is provided. The VLE it is extensively described and the IPA implementation is discussed as well, providing AI algorithms and IPA integration in SG.

- **Chapter 6** focuses on the assessment of the SG. In a first phase we have validated the IPA and tested it in three different classes involving roughly 50 high-school students. In a second phase we have tested the entire SG involving roughly 100 high-school students.
• Chapter 7 concludes and summarises this thesis, outlining future work and research directions.
Chapter 1

Introduction

Education and AI can be seen as two sides of the same coin: education helps students learn and extend the accumulated knowledge of a society and AI provides techniques to understand the mechanisms underlying thought and intelligent behaviour, Woolf et al. in [110]. In recent years there has been an increasing interest in the use of digital games for learning and behaviour change. Several of the studies demonstrate that playing games can change behaviour in a positive way. It seems also that “unintentional” learning found in entertainment games could provide insights into engagement and learning in Serious Games, Boyle et al. in [11].

The research described in this thesis focuses its attention on instructional systems involving Artificial Intelligence (AI) such as Intelligent Tutoring Systems (ITS). The purpose of such systems is to capture the very knowledge that allows experts to compose an instructional interaction in the first place. Instead of decisions resulting from some knowledge, it is the knowledge itself that is explicitly represented so that it can be used in computer based systems. It is then the responsibility of programs to compose instructional interaction dynamically, making decisions by reference to the knowledge with which they have been provided. The aim of this branch of AI is to explicit encode domains and pedagogical knowledge to have more intelligent, adaptive and effective behaviour in Tutoring Systems, Wenger in [106].

Since ITSs assume that the learner uses computerised support instead of the human instructor, Virtual Characters have been developed to improve interactivity, trying to compensate for the lack of human aspects. The problem of the lack of face-to-face interaction had consequences in motivational aspects, Soliman & Guetl in [99], but ITSs can support high degrees of individualization not possible in classical classrooms. Such tools were really successful in providing strong individualization to the learner by supporting adaptations to the system based on user models.

In the domain of e-learning scenarios, research has shown that highly interactive VLEs imprints a permanent transformation of the educational landscape, as it has the ability to produce better academic results, Aldrich in [2]. VLEs extend ITSs by
integrating several educational resources to the learner including multimedia learning material, communication tools, recommender systems, and more. In a VLE, learners are not restricted by time or space. 3D Immersive VLEs add 3D visualization and the ability of the learner to navigate the 3D virtual environment. The visualization of the learning space and interacting with other learners gives new possibilities of computer-aided learning scenarios not existed before in a game-like environment that opens the door for creativity and imagination to the learner.

Research has highlighted motivational factors from games, visualization and HCI aspects that can greatly motivate the nowadays learner. The emergence of game-based learning approaches has recently become popular with learners, tutors and policy makers because games can engage and motivate learners, including those learners who are disengaged from the traditional methods of chalk and talk and age and stage.1

This chapter provides an introduction to some of the pedagogical motivations and reasoning for introducing new technologies in teaching practice. In addition, it provides a link among Intelligent Tutoring System, Educational Games and Artificial Intelligence.

1.1 21st Century Skills

The Four Pillars identified by the UNESCO’s Task Force on education for the 21st Century deal with learning to know, learning to do, learning to live together, learning to be. Underpinning these pillars there are the 21st Century skills that have been very hard to define and model in traditional learning conditions, De Freitas & Ott in [35]. However, once defined, the task of teaching all these skills to a vast variety of students requires a major rethink of what education is and how it can be delivered to answer the 21st Century challenges.

This rethink in education follows the need to reshape and reconsider the requirements for new competences tailored for the “generation.com” or the new “millenials”, Garris et al. in [39]. In the past, education could have simply been considered as a way to deliver a framework for prescribed educational knowledge and curriculum. Today, the requirements of knowledge management have prescribed new sets of skills needed to adapt to fast changing technological advances. In light of these changes, e-learning methods have emerged, while a parallel change in society has driven interest in social communities, game play and use of web-based services into the forefront of teachers’ work. Leisure time and education are part of this transition of technology enhanced learning into classroom practice.

In favour of adopting highly interactive virtual environment, Aldrich in ([2]) presents the following three arguments:

- **Games as learning tools**: Games are a more natural way to learn than traditional classrooms. As they are the most ancient vehicle for education, they are able to create the optimal learning state and immerse the learners into specific contexts that build knowledge and skills.

---

1 Those two expressions are used to indicate traditional methods of teaching, focused on the blackboard and presentation by the teacher as opposed to more informal or interactive methods.
1.2 Artificial Intelligence grand challenges in education

- **Context and emotional Involvement**: Knowledge becomes entirely useful only in context. Serious Games can provide the context in which the educational content can be used and also bring the learner into an emotional state that stimulate memory.

- **Participation**: Participation with content may be necessary for learning. The process of converting experiential expertise into linear materials such as books might strip out what is the most valuable in the content.

In order to understand whether Educational Games are appropriate tools to contribute to build the so called 21st Century Skills it is necessary to clarify what these new skills are. As highlighted by Kickmeier-Rust & Albert in [54], the major obstacle to reaching a shared definition arises from “unclear, probably vague, and highly informal nature of these 21st Century skills”. The concept is then an overarching term for many kinds of meta-abilities, soft skills, communication and collaboration, self-awareness, strength in non linear thinking and innovative problem solving as well as the ability to reflect about one’s own thinking and being, Ott et al. [79].

It is important to highlight that Serious Games do not have the potential to develop such skills per se. The idea is that these skills can be addressed while developing a Serious Game and can be used as one of the lenses through which looking at the game.

The figure below, re-adapted from Ott et al. [79], has been used in this research project has a guide to address the so called 21st century skills.

![Diagram of 21st Century skills and Serious Games](image)

**Figure 1.1.** 21st Century skills and Serious Games

1.2 Artificial Intelligence grand challenges in education

“Artificial intelligence impacts growth and productivity in many industries (e.g., transportation, communication, commerce, and finance). However, one painful exception is education. Today, very few AI-based learning systems are consistently used in classrooms or homes. Yet the potential for an impact on education is great:
today’s instructional software now routinely tailors learning to individual needs, connects learners together, provides access to digital materials, supports decentralised learning tools and engages students in meaningful ways” suggests Woolf et al. in [110]. In response to the above criticism, Woolf introduces five grand challenges in education, which are:

- **Grand Challenge 1**: Mentors for every learner

- **Grand Challenge 2**: Learning 21st century skills

- **Grand Challenge 3**: Interaction data to support learning

- **Grand Challenge 4**: Universal access to global classrooms

- **Grand Challenge 5**: Lifelong and lifewide learning

These challenges are aligned with the goals of making learning more social, collaborative, inquiry-based, ubiquitous, accessible, pervasive, secure and available to people anytime and anywhere. They are intended to spur significant development in AI and highlight the richness of educational challenges.

In this thesis, we focus the attention on the first and second grand challenge: mentors for every learner and 21st century skills 1.1.

To mentor effectively and support individuals or groups while learning, an intelligent system needs to assess learning activities and model the changes that occur in learners. Estimates of a learner’s competence or emotional state, stored in user models, represent what learners know, feel, and can do. When and how was knowledge learned? What pedagogy worked best for this individual or group? Machine learning and data mining methods explore the unique types of data that derive from educational settings and use those methods to better understand students and the settings in which they learn, Conati & Zhao in [21]. AI provides the tools to build computational models of skills, learning processes, and scaffolding of learning. Further, AI methods can act as a catalyst for computer-based learning environments through the integration of cognitive and emotional modelling, knowledge representation and reasoning, natural language question answering and machine learning methods, into software to provide knowledge about the domain, student and teaching strategies, Woolf in [109]. Such systems provide flexible and adaptive feedback to students, enabling content to be customised to fit personal needs and abilities and augment a teacher’s ability to respond. These are essential ingredients for achieving the vision of mentors for every learner and represent both ongoing and future areas of AI research. Providing a mentor for every learner and groups of learners means improving the ability of systems to provide timely and appropriate guidance. In other words, the determination of what to say to learners, when to say it, and how to say it grows more complicated as the skills demanded by society also increase in complexity. The learning sciences has provided a wealth of knowledge about how to deliver effective feedback, but the aim to incorporate 21st century skills, such as creativity and teamwork, present new challenges. Rich, multi-faceted models of instruction and coaching will be needed that go beyond simple hinting: these systems must leverage advances in new user modelling techniques and the richness of modern learning environments that are increasingly more social and immersive.
Further, AI-based systems are emerging that focus on affective issues, such as emotional self-regulation and behaviour change. These require reconsideration of the role of feedback and more robust systems that seek to balance the cognitive aspects of learning with the non-cognitive. Future learning environments should seek to build confidence in learners, inspire interest in important topics like science, promote deep engagement in learning, and reduce or eliminate the barriers to learning present in the world today. In recent years, there has been an increasing interest in how to make computers more “sociable” by enabling them to both display their own emotions and react to the user’s emotions. Building computers that display emotions in a natural and meaningful way is already a challenging endeavour, since it requires formalizing concepts and mechanisms that are often still under investigation in emotional psychology. Building computers that recognise user’s emotions is even more challenging, as proven by the fact that even human beings are not always proficient in this task. The challenge is due to the high level of ambiguity that exists in the mapping between emotional states and the factors that can be used to detect them. For instance, different people can have different emotional reactions to the same stimulus, and the variability depends upon traits that are not always easily observable, such as a person’s goals, preferences, expectations, and personality as explained by Conati in [20]. Emotions can be recognised because they often have observable effects on a user’s behaviour and bodily expressions, but the mapping between emotions and their observable effects also depends on often hidden traits of a person, as well as on the context of the interaction.

1.3 Game Based Learning

In line with what has been claimed so far, in recent years there has been an increased interest in the use of digital games for learning and behaviour change with an increasing number of conferences, journals, projects and societies devoted to the topic. The term game is quite ambiguous, in the sense that researchers, game designers, parents, students, teachers, etc have a different concept of games. This is because games refer to a lot of different game formats (video games, location-based games, board games, etc.), game genres (strategy games, edutainment game, role playing game, etc.) and game dynamics (competition, exploration, resource management, etc.). To avoid confusion about games in a learning context, Game Based Learning (or teaching) is often combined (or used synonymously) with the term Serious Games. Hereafter, we provide some definitions of Serious Game and Game-Enhanced Learning, proposed by Dyer et al. in [93]:

- According to Corti\(^\text{1}\) a Serious Game “is all about leveraging the power of computer games to captivate and engage end-users for a specific purpose, such as to develop new knowledge and skills”
- In [114], Zyda defines serious games as “a mental contest, played with a computer in accordance with specific rules, that use entertainment to further government or management education, health, public policy and strategic communication objectives”
- In [70], Micheal & Chen define a serious game as “a game in which education (in its various forms) is the primary goal, rather than entertainment”
In the context of this research study we have adopted the definition of SGs as inquiry based laboratories in which participants are able to imagine, engage with, and reflect upon their experiences. Games are intended as scenarios, following The definition given by Hanghøj in [93]. Scenarios directly refer to the dynamic, future oriented models for possible actions that are embedded into game designs. Games are well suited for developing students’ scenario competence, which can be defined as the ability to imagine, enact and reflect upon game-specific choices and their consequences. Thus scenario competence represents a form of meta-competence, which involves three different aspects: problem scenario, social scenarios and identity scenarios. These three aspects correspond to the three major functions of education: qualification, socialization, and subjectification, as explained by Biesta in [7]. Unfortunately, gaming and schooling have developed into two distinct “knowledge traditions” that often rely on opposing validity criteria for determining what counts and what does not count as relevant knowledge. To avoid that dichotomy, Serious Games should integrate different aspects that are related to the knowledge itself, to pedagogical aspect, to scenario-based and every day practise. Hanghøj suggests that there is a complex translation involved in using games for educational purposes and thus Serious Games can be seen as a dynamic interplay of four knowledge practices, as depicted in Figure 1.2.

![Figure 1.2. Interplay of the four knowledge practices](image)

Much of the interest in SGs has been speculative, discussing the potential of games to provide new methods for supporting learning. To address concerns about the lack of empirical evidence about the effectiveness of games, Boyle et al. [11] carried out a literature review that aimed to identify research evidence about the positive impacts of games. The review advocated a narrative approach to capture the diversity of the selected studies (empirical papers published in the five year period since the end of the previous review, i.e. between March 2009 and February 2014) and developed a multi-component analysis of games and their outcomes which provided a useful framework for organising the research along key variables, making...

---

links between the different ways of categorizing games and outcomes, and identifying emerging trends, hot topics and gaps in the literature. The review done by Boyle et al. recognises the heterogeneity of the research in this relatively new field and aims to provide some organisation and structure to this area. Using the same search terms on a smaller number of databases, the review found many more papers reporting empirical evidence of the positive outcomes of playing games (512) than the previous review (129), carried out in 2009. This illustrates the increased interest in the positive impacts of digital games during the five year period from 2009 to 2014 compared to the previous five year period. Given the very large number of relevant papers identified, the review focused on summarising the 143 papers that were rated as higher quality. Extending the multidimensional classification to consider the geographical location of the first author confirmed that research on games is being carried out across the globe, especially in North America and Europe. Coding the papers in terms of the general aim of each paper confirmed that, researchers recognise the importance of studying player characteristics, game features and the context of play as variables that influence game outcomes. This analysis is consistent with models and theories of game design, such as De Freitas et al. in [24] and Peachey in [81], that emphasise the need to consider the coordination of a range of components in designing and evaluating games.

1.4 Combining technical and pedagogical challenges

The aim of the research presented in this report is to introduce a new SG, completely designed and realised coping with the technological and pedagogical challenges discussed so far. The need to realise from scratch a new SG derives from different boosts. In fact, although research on SGs have been very prolific in current and past years, major efforts have been made in games targeted at primary and middle school. Another area in which SG are well explored is for training purposes with adults or at university level. Unfortunately there are very few SGs targeted at high school students, which is the target of our SG. Many SG come from researchers and very often they remain just in a prototype phase so it is difficult to reuse or test them in real case scenarios. Finally, from a technological perspective, very often SG are concentrated on VLE or on ITS, it is very uncommon to find a SG that combine those two aspects, Soliman & Guetl in [99]. The research’s challenges addressed in this research are shown in Figure 1.3.

Specific pedagogy and psychology issue plus 21st century skills have been used as a guide for the design of the Serious Games. From a pedagogy point of view, beside the 21st century skills, there are many more factors that influenced the choices made in the design and development of the SG, as deeply explained in Chapter 3.
In summary, we can say that research in psychology and pedagogy field combined with the necessity to address 21st century skills lead us to build a new Serious Game that make use of IPA and VLE. The main novelty of this research thesis, from an engineering point of view, has not been to improve all the aspects listed in Figure 1.3 but to give an original contribution on some specific issues, as explained in the following box.

**Research Advancement**

The main original contributions reported in this thesis have been two. First, the design and implementation of an Emotional Intelligent Pedagogical Agent (EIPA) as a response to the AI grand challenge number one (mentors for every learner, Woolf et al. in [109]) integrated in a Virtual Learning Environment that responded to all the requirements traced by Dillenbourg in [29]. Second, the development of an AI algorithm that provides unsolicited hints during game session taking into account games’ variables and players’ emotions, which is a rather unexplored field.
Chapter 2

Virtual Learning Environment and Intelligent Tutoring System

In the domain of e-learning scenarios, research has shown that highly interactive VLEs imprints a permanent transformation of the educational landscape, as they have the ability to produce better academic results, Aldrich in \[2\]. Since VLEs assume that the learner uses computerised support instead of the human instructor, Virtual Characters have been developed to improve interactivity, trying to compensate for the lack of the human aspects in ITS, as explored by Soliman & Guetl in \[99\].

In this chapter we introduce the two main topics that have been tackled in this research, that are Virtual Learning Environments and Intelligent Tutoring Systems.

2.1 Virtual Learning Environments

Before going deep into the review of some implemented VLEs we provide here a definition of VLE that is commonly accepted and make some clarification about the term. In fact sometimes the term VLE is over abused or used in a wrong way. According to Dillenbourg, as explained in \[29\], a Virtual Learning Environment must comply with some specific rules:

1. A virtual learning environment is a designed information space. For learning environments, the functional requirements are numerous and have not yet been systematically studied. However, here are a few examples of functional requirement that a VLE must satisfy:
   
   • using information in educational interactions (information must be dynamically accessed and recovered).
2.1 Virtual Learning Environments

- multi-authoring (there must be some mechanism that allows different authors to collaborate);
- indicating information source;
- maintaining information;
- following technical evolution;
- sharing information with the world.

2. A virtual learning environment is a social space. What is specific to virtual environments compared to any information space is that it is populated by the information itself. The users are inside the information space and see a representation of themselves and/or others, in the space. As soon as students see who else is interested by which information, the space becomes inherently social.

3. The virtual space is explicitly represented. The representation of the learning environment ranges from text-based interfaces to the most complex 3D graphical output. The key issue is not the representation per se, but what the students actually do with this representation.

4. Students are not only active, but also actors. The notion of a learning activity in virtual learning environments refers to something richer than in individual courseware, closer to the notion of project. The difference between other constructivist environments and what virtual environments potentially offer can be described as making students not only active, but also actors, i.e. members and contributors of the social and information space.

5. Virtual learning environments are not restricted to distance education. The use of VLE can influence the way teachers teach and thereby contribute to renew teaching methods in classroom practice.

6. Virtual learning environments integrate multiple tools. A physical learning environment generally integrates courses, resources (libraries), formal communication (boards) and informal communication (cafeteria, ...), an administration, etc. Similarly, a virtual learning environment integrates a variety of tools supporting multiple functions: information, communication, collaboration, learning and management. The idea of environment includes this notion of integration.

7. Most virtual environments overlap with physical environments. Virtual learning environments do not only integrate a variety of software tools but also integrate all the physical tools that can be found in a classroom.

Bearing in mind the 7 rules traced by Dillenbourg VLE has been clearly defined but can still be developed in many different ways. In fact a VLE can be mainly web based (such as e-learning platform) or an immersive and navigable 3D environment (like many games). Three-dimensional technologies have become a fundamental element of almost all modern computer games, including most of the current massively multi-player online
2.1 Virtual Learning Environments

games (MMOGs) such as World of Warcraft, which is the most popular MMORPG game in terms of the number of online subscribers with 10 million subscribers as of January 2008. 3D Virtual Environment (VE) are also central to the new generation of immersive virtual worlds, such as Active Worlds and Second Life (SL). A 3D VE can be defined as an “environment that capitalises upon natural aspects of human perception by extending visual information in three spatial dimensions”, “may supplement this information with other stimuli and temporal changes and enables the user to interact with the displayed data”, following Wann & Mon-Williams in [103]. Three dimensionality, smooth temporal changes and interactivity are the most important features that distinguish 3D VLEs from other types of VLEs, such as those provided by a learning management system (e.g., Blackboard or Moodle).

Many authors have stressed the importance of immersion and presence (i.e. Mikropoulous in [71] and Jonassen in [51]), suggesting that these are critical features distinguishing (VEs) from other types of computer applications. In early writings about VEs, there was a tendency to use these terms interchangeably; subsequently, debates occurred in the literature about the definitions of these terms, who defines presence as the subjective sense of being in a place, and immersion as the objective and measurable properties of the system or environment that lead to a sense of presence. In other words, immersion relies on the technical capabilities of VR technology to render sensory stimuli, whereas presence is context dependent and draws on the individual’s subjective psychological response to VR. The latter is dependent on a range of factors including, but not limited to, the user’s state of mind, Slater in [97].

Dalgarno & Lee in [23] adopt the perspective that representational fidelity and learner interaction are unique characteristics of 3D VLEs, whereas construction of identity, sense of presence and co-presence are characteristics of the learner’s experience as a result of these environment characteristics. They argue that the degree that 3D VLEs have the potential to provide learning advantages over non-3D resources, in particular, is dependent on a number of underlying, general assumptions about cognition and learning in 3D environments, along with assumptions about links or connections between the distinguishing characteristics of 3D VLEs and the potential or anticipated learning benefits. The overall picture have been re-elaborated in Figure 2.1.

For the sake of simplicity, the applications reviewed in the next paragraph have been grouped into three broad categories: 3D simulations and microworlds, 3D environments as interfaces to learning resources, and 3D multi-user VLEs.

### 2.1.1 3D simulations and microworlds

Simulations have been used as part of computer-assisted learning materials for at least three decades, with SimCity (Wright in [111]) being one of the earliest and most popular examples. Simulated 3D environments modelled on real places and objects have the potential to provide an enhanced sense of realism and a greater sense of presence as compared with non-3D environments. Their fidelity is such that where barriers exist for visiting the real place, immersion in the 3D VE can be a viable alternative. For example, Alberti et al. [1] describe a 3D VE modelled on a historic theatre in Italy, while Kontogeorgiu et al. in [56] describe the exploration of 3D
simulated microscopic environments in an effort to allow students to experience being inside a quantum atom. In Virtual Big Beef Creek by Campbell et al. [13], a 3D VE that recreates a marine and coastal environment to assist in the teaching of ocean science, learners can assume the roles of scientists to collect and analyse geo-scientific data or alternatively, take on characters representing creatures that inhabit the environment, which are variously able to walk over land, swim underwater or fly across the sky. In this way, the learners are able to explore first-hand the abilities and limitations of the various animals while simultaneously acquiring knowledge about the flora, fauna, ecosystem and ocean environment at large. Other examples of 3D microworlds for learning are portrayed by Kaufmann et al. in [52] and Yeh & Nason in [112], who describe 3D environments for developing learners’ understanding of geometry; the simulated radioactivity laboratory described by Crosier et al. in [22] allow learners to carry out tasks and measure the results at the laboratory level, then zoom in and visualise what is happening at the atomic level. Furthermore, 3D microworlds may be used to allow the learner to construct his or her own 3D environment and/or objects as a way of articulating his or her spatial model or externalising his or her understanding of particular abstract concepts, Winn in [107]. Many 3D concept mapping tools have been developed for such purposes, examples of which are Nelements1 and Topicscape2.

2.1.2 3D environments as interfaces to learning resources

Many studies have found that learners can have difficulty navigating hypermedia environments, with the problems characterised by the lost in hyperspace phenomenon, as suggested by McKnight et al. in [68], whereby users lose track of how they arrived at a node and have no clear model of the overall environment structure. The

---

1http://nelements.org/, accessed January 12, 2017
2https://www.topicscape.com/, accessed January 12, 2017
provision of an interface that allows easy navigation through the information, while maintaining a sense for the overall structure of the resources and the connections between ideas, is problematic. Three-dimensional environments offer transparency of knowledge representation, which allows learners to approach concepts as first-person, non-symbolic experiences, in contrast to most instances in which information is codified and represented as third-person, symbolic experiences, Winn in [107]. These applications attempt to capitalise on learners' well-developed spatial cognitive abilities to assist them in navigating within the information space, as explained by Liang & Sedig in [62]. In [14], Card et al. propose the use of a 3D environment as an interface for navigating through a complex information space and in [86], Robertson et al. description of the use of a 3D interface for task management on a PC are consistent with this idea. Applications such as these provide examples of the ability of 3D VLEs to facilitate learning tasks leading to spatial knowledge representation as early depicted in Figure 2.1. The formation of a spatial cognitive model of the information space as a result of exploring an environment has the potential to boost exploration efficiency and conceptual understanding of the learning domain.

2.1.3 3D multi-user VLEs

In [26], Dede et al. discussed the possibility of combining the capabilities of VEs with the capabilities of computer-mediated communication tools to promote collaborative learning within a distributed virtual space. Today’s multi-user, distributed 3D environments, including MMOGs and virtual worlds, allow geographically dispersed users to explore an environment concurrently, with each represented by a surrogate persona or avatar visible to other users and with tools allowing text-based or audio communication teacher or facilitator as a learner undertakes learning tasks. However, they also have great potential as social learning and computer-supported collaborative learning tools, according to Edirisingha et al. in [31]. According to social constructivist views of learning, conversation and discourse are the cornerstones of collaboration and social negotiation in learning, Jonassen in [50]. Communication within a simulated environment relevant to the ideas being discussed can provide a greater sense of place than other text-based alternatives such as instant messaging, chat rooms and multi-user dungeons/dimensions (perhaps better known as MUDs), and consequently, it can help foster greater closeness within the group and richer communication because of the ability to draw on spatial and non-verbal cues. If role-play strategies are used, it is likely that learners will easily lose themselves as they adopt their role and identify with their avatar because of the fidelity of the environment. This ability to self-define and take on alternate personae gives learners opportunities to adopt multiple perspectives, the importance of which is a key tenet of constructivist learning, according to Honebein in [45], while the willed suspension of disbelief and emotional realism encourage them to engage in exploration, inquiry and risk taking, as explained by Dickey in [27]. In [69], Mennecke et al. report on how students undertake a scavenger hunt activity in Second Life, in which they co-experience and explore the virtual world as they embark on a mission to discover interesting places and practise basic SL skills. To complete the exercise, they must retrieve the relevant instructions, decipher the embedded hints and teleport to the location of the item they are searching for. The activity
requires students to work in teams, communicating and coordinating their activities and collaborating in the process. Successful completion is achieved when the team leader submits a note card containing details of the team’s collaboration as outlined in the scavenger hunt instructions. In another example [49], Jarmon et al. tell of how students in a graduate-level communication course work together and in collaboration with architecture students at the same university. The communication and architecture students are tasked with creating a virtual presence in SL of two green, sustainable, urban housing designs, that are later physically implemented in a low-income neighbourhood in Austin, Texas. Successful completion of the course assignments and projects is contingent on the students in both disciplines interacting extensively with educational and non-academic participants, both in real life and in the 3D virtual world. Positive interdependence is also evident in that the communication students are reliant upon the domain knowledge and expertise of the architects, and vice versa.

2.2 Intelligent Tutoring Systems

We define an ITS as a computer learning environment that helps the student master deep knowledge/skills by implementing powerful intelligent algorithms that adapt to the learner at a fine grained level and that instantiate complex principles of learning, following Graesser et al. in [43]. We can see ITS environments as a generation beyond conventional Computer Based Training (CBT). CBT systems also adapt to individual learners, but they do so at a more coarse-grained level with simple learning principles. In a prototypical CBT system, the learner:

- studies material presented in a lesson
- gets tested with a multiple-choice test or another objective test
- gets feedback on the test performance
- re-studies the material if the performance is below a certain threshold
- progresses to a new topic if performance exceeds threshold

The order of topics presented and tested typically follows a predetermined order, such as complexity (simple to complex) or prerequisites. The materials in a lesson can vary from organised text with figures, tables, diagrams and multimedia, to example problems to be solved. ITSs can be viewed as enhancements of CBT with respect to the adaptability, grain-size, and the power of computerised learning environments, O’Neil et al. in [78].

In ITSs, the processes of tracking knowledge (called user modelling) and adaptively responding to the learner incorporate computational models in Artificial Intelligence and Cognitive Science, such as production systems, case-based reasoning, Bayes networks, theorem proving, and constraint satisfaction algorithms, Graesser et al. in [43]. Successful ITSs have been developed for mathematically well formed topics, including algebra, geometry, programming languages Aleven et al in [3], physics Freedman in [34] and information technology, Mitrovic et al. in [72]. Some
2.2 Intelligent Tutoring Systems

intelligent systems handle knowledge domains that have a stronger verbal foundation as opposed to mathematics and precise analytical reasoning. Auto Tutor, as described by Graesser et al. in [42], helps college students learn about computer literacy, physics, and critical thinking skills by holding conversations in natural language. Other natural language ITSs, that have shown learning gains, include DeepTutor, iSTART, and My Science Tutor, described respectively in the work by Rus et al. [89], Jackson et al. [47] and Ward et al. [104].

ITS technologies that target deep learning have the challenge of keeping students motivated because there is an inverse relationship between liking a system and deep learning. Research suggest that games can fill the motivational gap for ITSs as explained by Soliman & Guetl in [99], and O’Neil et al. in [78]. However, there have not been enough studies combining games with ITSs for a meta-analysis of the outcome under the afore mentioned lenses. However, some successful games that integrates ITS and VT have been developed so far with encouraging results.

Crystal Island, Rowe et al in [88], is a narrative-centered learning environment built on the Valve Software’s Source™ engine, the 3D game platform for Half-Life 2. Crystal Island features a science mystery set on a recently discovered volcanic island. The curriculum underlying Crystal Island’s science mystery is derived from the North Carolina state standard course of study for eighth-grade microbiology. Students play as the protagonist, Alyx, who is attempting to discover the identity and source of an infectious disease plaguing a newly established research station. The story opens by introducing the student to the island and members of the research team for which the protagonist’s father serves as the lead scientist. Several of the team’s members have fallen gravely ill, including Alyx’s father. Tensions have run high on the island, and prior to Alyx’s arrival various team members began to accuse one another of having poisoned the sick researchers. It is the student’s task to discover the outbreak’s cause and source, and determine whether one of the team members is guilty of poisoning. Narrative-centered learning environments are inherently complex systems, and multiple lines of investigation are necessary to develop suites of technologies that can produce adaptive, engaging learning experiences. At the time of publication, three principal areas of research have been conducted with the Crystal Island virtual environment: narrative generation and reasoning, affect recognition and expression, and empirical evaluation of NLEs. The conclusions are that narrative is the subject of increasing attention in the AI in Education community as a powerful medium for contextualizing learning. Narrative-centered learning environments present a range of opportunities for investigating how different computational models can be leveraged to create effective, engaging learning experiences. Work on the Crystal Island environment has begun to illustrate how models of narrative generation and reasoning, as well as affect recognition and expression, can bear on game-based learning environments. A series of empirical evaluations has begun to demonstrate the motivational and pedagogical potential of narrative-centered learning.

The interactive Strategy Training for Active Reading and Thinking (iSTART-ME) tutor, Jackson et al. in [48], is a web-based reading strategy trainer that provides young adolescent to college-aged students with reading strategy training to better understand challenging science texts. iSTART responds to the need of reading comprehension interventions, by providing reading strategy instruction in an automated web-based system. In iSTART, pedagogical agents instruct trainees
in the use of self-explanation and other active reading strategies to explain the meaning of science text while they read. The training was motivated by empirical findings that show that students who self-explain text are more successful at solving problems, more likely to generate inferences, construct more coherent mental models, and develop a deeper understanding of the concepts covered in the text. The results from this study indicated that learning is better for the non-game environment, but that engagement is better for the game-based system. The differences in learning may largely be due to the different pedagogical approaches employed within the respective training environments, as explained by the authors. One module includes fine-grained feedback designed to improve the quality of subsequent self-explanations, and the other implements modelling by presenting example self-explanations from an opponent player. Performance results suggest that the fine-grained feedback provides better support during the application of the self-explanation strategies. In contrast to the performance differences, students expressed increased engagement and persistence within the module with the opponent player. While coached practice may have produced higher quality self-explanations during the experiments, the authors claim clearly that students preferred the interaction with the game-mode.

Despite the examples made so far, there is no adequate body of research reporting effect sizes that contrast the game versions versus those without game features in these ITSs.

In this thesis, we have focused our attention on a sub branch of ITS, mainly on system capable of making a conversation in natural language: namely Conversational Agents (CAs). As pointed in one of the innovation of the developed SG is to have an Intelligent Pedagogical Agent (facing the first AI grand challenge in education mentioned in 1.2) inside a Virtual Learning Environment. Since the scope of the Virtual Tutor that we imagined is to assist the learner during the game evolution, we decided to implement it as a Conversational Agent, capable of pedagogical attitudes, as explained in Sections 2.2.1 and 2.2.2.

### 2.2.1 Conversational Agent

The use of interaction mechanisms based on natural language represents an interesting feature for applied games and a valid opportunity to improve aspects related to the communication and the transfer of knowledge. Considering the application of natural language interaction techniques, chatbots are conversational agents that use NLP techniques and are often employed in Web sites as virtual assistants. Their knowledge is defined though syntactic rules and scripting languages specifically designed for this purpose. Conversational ITSs have several advantages over other types of ITSs. They encourage deep learning as students are required to explain their reasoning and reflect on their basic approach to solving a problem. Conceptual reasoning is more challenging and beneficial than mechanical application of mathematical formulas. Furthermore, conversational ITSs have the potential of giving students the opportunity to learn the language of scientists, an important goal in science literacy. A student associated with a more shallow understanding of a science topic uses more informal language as opposed to more scientific accounts, Rus et al. in [89]. The impact of conversational ITSs allegedly can be augmented by the use of animated conversational agents that have become more popular in contemporary advanced
learning environments, Graesser in [41]. The animated agents interact with students and help them learn by either modelling good pedagogy or by holding a conversation with the learners. Among the most popular chatbot technologies, we can mention ALICE (Artificial Linguistic Internet Computer Entity) and ChatScript. ALICE technology, see Wallace in [102], is based on AIML (Artificial Intelligence Markup Language), which is an XML-like mark-up language. ChatScript is a scripting language upon which several award-winning bots, like Suzette and Rozette, are based. An interesting example of an AIML-based chatbot applied to a cultural context is the work by Santangelo et al. [92]. This work concerns the implementation on mobile devices of a chatbot for a visit to a cultural heritage site. The dialogue is managed by exploiting a knowledge base described by question-answer modules. However, these modules are structured with AIML, and are not described by any authoring tool. In order to give an idea of how complex the definition of the knowledge usually is using a standard rules-based system, we can mention a talk\(^1\) given by the developer of ChatScript, Wilcox, who reports that an ALICE bot uses about 120,000 rules, while Suzette and the most recent Rozette, respectively need about 15,000 and 10,000 rules.

It becomes clear that asking a user to specifically learn a scripting language to program such a high number of rules to define the knowledge for just one character becomes infeasible. It is essential, in our research, to leverage on a system that allows the author of the knowledge to not be constrained by the use of complex rule-based systems and that makes the authoring process as simple as possible. Chatbots, however, are typically employed as chatting agents and their purpose is to give the illusion of intelligence, our goal instead refers to define character who converse on specific topics and not on general subjects. Populating applied games with conversational agents is not a novelty, they have been previously employed in order to create alive virtual environments, where the user can feel immersed in the action, which takes place in a realistic environment, thus creating the best conditions for effective learning. Some examples in this context are the work done by Oberlander et al. in [76] and Chittaro et al. in [17], which employ characters to guide the user in a virtual archaeological museum set in Second Life and in a virtual reconstruction of a real museum. The interactions in these examples are only passive, since the user is only allowed to receive information from the guide, without asking questions. Finally, Neto et al. in [74] implements interactive storytelling through an embodied agent, who from time to time during gameplay, asks the user a question to verify his learning status. The user can interact by speaking to the system, however the narration is scripted, and the user capabilities are limited to the choice of one among four predetermined answers. Other examples of successful conversational ITSs beside AutoTutor, are Why2 (VanLehn et al. in [101]) and GuruTutor (Olney et al. in [77]).

The VT that we have developed advance over previous examples by providing as-a-service the AI algorithm. The advantage of providing the VT has a service is mainly due to the fact that it will be easy to reuse it in any part of the game. In addition, it could reside on the server side, having the client more light as better

\(^1\)http://www.gamasutra.com/blogs/BruceWilcox/20120104/90857/Fresh_Perspectives_A_Google_talk_on_Natural_Language_Processing.php, accessed January 17, 2017
explained in chapter 5. We defines a bidirectional system that allows the player to express his own thoughts directly to the character. The character, in turn, must be programmed in order to comprehend the natural language sentences expressed by the user.

### 2.2.2 Intelligent Pedagogical Agent

ITSs can deal with emotional learning as well. Emotions are important for students in two major ways. First, emotions have an impact on learning. They influence our ability to process information and to accurately understand what we encounter. For these reasons, it is important for teachers to create a positive, emotionally safe environment to provide for the optimal learning of students. Then, learning how to manage feelings and relationships constitutes a kind of *emotional intelligence* that enables people to be successful. Emotional intelligence expands on Howard Gardner’s [19] theory of multiple intelligences, in particular, the intra-personal and inter-personal intelligences he defines, which deal with understanding oneself and others.

One of the most influential papers in the area of emotional learning is a study Lester at al. in [61], dated 1997, that demonstrated a Persona Effect, in which learning was facilitated by an animated pedagogical agent that had a life-like persona and expressed affect. The rationale for this research has been the media equation hypothesis of Reeves & Naas in [85]. They suggest that learners respond to pedagogical agents as if they were social actors, then the agents’ effectiveness should depend upon whether or not they behave like social actors. The agents should be socially intelligent, acting in a manner that is consistent with their social role, in accordance with social norms. In fact, human tutors make extensive use of social intelligence when they motivate and support learners, Lepper et al. in [60]. Thus, social intelligence in pedagogical agents may be important not just to gain user acceptance, but also to increase the effectiveness of the agent as a pedagogical intervention.

We list hereafter the most common approaches to the study of emotions and the difficulties that arose trying to automatically recognise emotions in HCI. In fact, while automatic detection of the six basic emotions in posed controlled audio or visual displays can be done with reasonably high accuracy, detecting these expressions or any expression of human affective behaviour in less constrained settings is still a very challenging problem due to the fact that deliberate behaviour differs in visual appearance, audio profile, and timing from spontaneously occurring behaviour.

**Database:** Having enough labelled data of human affective expressions is a prerequisite in designing an automatic affect recogniser. Most of the existing studies on the automatic analysis of human affective displays have been based on the artificial material of deliberately expressed emotions, elicited by asking the subjects to perform a series of emotional expressions in front of a camera and/or the microphone. These findings and the general lack of a comprehensive reference set of audio and/or visual recordings of human affective displays motivated several efforts aimed at the development of data sets that could be used for training and test of automatic systems for human affect analysis. In most of the existing databases, discrete emotion categories are used as the emotion descriptors. The labels of prototypical emotions
2.2 Intelligent Tutoring Systems

are often used, especially in the databases of deliberate affective behaviour. In databases of spontaneous affective behaviour, coarse affective states like positive versus negative, dimensional descriptions in the evaluation-activation space, and some application-dependent affective states are usually used as the data labels. The existence of these data sets of spontaneous affective behaviour is very promising, and according to researchers this will produce a major shift in the course of the research in the field: from the analysis of exaggerated expressions of basic emotions to the analysis of naturalistic affective behaviour.

**Vision-Based Affect Recognition**: Because of the importance of face in emotion expression and perception, most of the vision-based affect recognition studies focus on facial expression analysis. It is possible to distinguish two main streams in the current research on the machine analysis of facial expressions: the recognition of affect and the recognition of facial muscle action (facial AUs). Facial AUs are a relatively objective description of facial signals and can be mapped to the emotion categories based on a high-level mapping such as EMFACS and FACS/ID or to any other set of high-order interpretation categories, including complex affective states like depression or pain. Most of the existing facial expression recognisers employ various pattern recognition approaches and are based on 2D spatio-temporal facial features. The usually extracted facial features are either geometric features such as the shapes of the facial components (eyes, mouth, etc.) and the location of facial salient points (corners of the eyes, mouth, etc.) or appearance features representing the facial texture, including wrinkles, bulges, and furrows. Most of the existing 2D-feature-based methods are suitable for the analysis of facial expressions under a small range of head motions. Thus, most of these methods focus on the recognition of facial expressions in near-frontal-view recordings.

**Audio-Based Affect Recognition**: Research on vocal affect recognition is also largely influenced by a basic emotion theory. In turn, most of the existing efforts in this direction aim at the recognition of a subset of basic emotions from speech signals. In addition, few efforts toward the automatic recognition of non-linguistic vocalizations like laughter, coughs and cries have been reported recently. This is of particular importance for the research on the machine analysis of human affects since recent studies in cognitive sciences showed that listeners seem to be rather accurate in decoding some non-basic affective states such as distress, anxiety, boredom, and sexual interest from non-linguistic vocalizations like laughs, cries, sighs, and yawns. Although the above studies indicated recognition improvement by using information on language, discourse, and context, the automatic extraction of these related features is a difficult problem. First, existing automatic speech recognition (ASR) systems cannot reliably recognise the verbal content of emotional speech. Second, extracting semantic discourse information is even more challenging. Most of these features are typically extracted manually or directly from transcripts.

**Audiovisual Affect Recognition**: The data fusion strategies utilised in the current studies on audiovisual affect recognition are feature-level, decision level, or model-level fusion. Typical examples of feature level fusion are those which concatenated the prosodic features and facial features to construct joint feature vectors, which are then used to build an affect recogniser. However, the different time scales and metric levels of features coming from different modalities, as well as increasing feature-vector dimensions influence the performance of an affect recogniser.
based on a feature-level fusion. The vast majority of studies on bimodal affect recognition reported on decision level data fusion. In the decision-level data fusion, the input coming from each modality is modelled independently, and these single-modal recognition results are combined in the end. Since humans display audio and visual expressions in a complementary redundant manner, the assumption of conditional independence between audio and visual data streams in decision-level fusion is incorrect and results in the loss of information of mutual correlation between the two modalities. A number of model-level fusion methods have been proposed to address this problem, which aim at making use of the correlation between audio and visual data streams and relaxing the requirement of synchronization of these streams.

In [95], Shen et al. describe the development of an affective e-Learning model, and demonstrates the machine’s ability to recognise learner emotions from physiological signals. Pervasive e-Learning Platform aims to support pervasive learning environments where learning resources and tools could be accessed by students any time anywhere. The goal of this study was to understand how learners’ emotions evolve during learning process, so as to develop learning systems that recognise and respond appropriately to their emotional change. The proposed affective model, based on Bayesian Network study, indicated that the user’s emotional states were related to learner profiles (e.g., learning preferences, cognitive skills, knowledge structure), goals of learning, and learner interaction with the learning system. The authors claim that there is a lack of research efforts in detecting emotions during learning in real-time, the emotion evolution during learning and using the emotional feedback to better design e-Learning system. Findings of this study should make a contribution in the research and development of affective e-Learning systems. Conati in [20] presents an approach to a possible modelling of user affect designed to assess a variety of emotional states during interactions: knowing the details of a user’s emotional reaction can enhance a system capability to interact with the user effectively. This paper starts from some theoretical base and then explicitly encodes and processes the uncertainty by relying on probabilistic reasoning. In particular, they use Dynamic Decision Networks [25] to represent in a unifying framework the probabilistic dependencies between possible causes and emotional states (including the temporal evolution of these states), and between emotional states and the user bodily expressions they can affect, like the paper from Shen et al. mentioned above.

Zakharov et al. in [113] present a pedagogical agent capable of act by the logic which integrates the learner’s cognitive and affective states. They developed an algorithm for feature tracking which utilises a combination of common image processing techniques, such as thresholding, integral projections, contour-tracing and Haar object classification. Throughout the algorithm, the focus of attention is shifted among a number of regions of interest, determined on the basis of the anthropomorphic constraints describing human face geometry. The algorithm relies on a few session-dependent threshold values for processing eye, brow and mouth regions. The experiment results, explained by the authors, indicate a range of preferences associated with pedagogical agents and affective communication. Affective interaction is individually driven, and it is reasonable to suggest that in task-oriented environments affective communication carries less importance for certain learners. The authors conclude that, although the interplay of affective and cognitive processes
always underpins learning outcomes, affective interaction sometimes may need to remain in the background; whatever the case, an ITS should let the user decide on the level of affective feedback, if any, thus leaving the user in control.

**Figure 2.2. Emotion and Learning state of the art**

![Diagram](image)

IPA in Serious Games

Beside the possibility to have a dialogue with an Intelligent Virtual Tutor, it is very challenging to create educational games that are more pedagogically effective by making them capable of providing interactions tailored to each student’s needs and targeted at stimulating learning when necessary. Although there is well established research on building student-adaptive computer based educational tools, to date very little of this research has focused on electronic educational games. There is very few literature in this field. Following Conati in [20], there are two main challenges for this approach. The first one is that in educational games it is especially difficult to assess students’ knowledge and learning from the interaction with the game, because often game actions do not have a direct connection with a student’s understanding of the underlying domain. The second challenge is how to provide individualised interventions that trigger learning, without interfering with the high level of engagement that educational games usually trigger precisely because they do not remind students of traditional educational activities. Conati proposes a model in the context of the interaction with pedagogical agents designed to improve the effectiveness of computer-based educational games. The authors also introduce Dynamic Decision Networks and illustrate how they can be used to enable pedagogical agents for educational games to generate interactions tailored to both the user’s learning and emotional state. They describes in detail the DDN underlying their model of user affect and how it integrates, in a principle way, different sources of ambiguous information on the user’s emotional state. To make

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector, Lienhart & Maydt in [63].
the agent pedagogically intelligent in the game, it provides unsolicited hints to the user, analysing his behaviour during the game session.

In our research, in order to pass from an Intelligent Virtual Tutor to an Intelligent Pedagogical Agent we have followed the idea of the agent capable of providing unsolicited hints. In Table 2.1, we gave a specific definition of IVT and IPA, giving the meaning that those terms have taken in this specific research.

Table 2.1. IVT vs IPA: Definitions

<table>
<thead>
<tr>
<th>Intelligent Virtual Tutor</th>
<th>Intelligent Pedagogical Agent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Virtual Agent capable of a conversation in natural language</td>
<td>A Virtual Agent capable of a conversation in natural language and of actuating intervention in the game to support specific user needs</td>
</tr>
</tbody>
</table>
Chapter 3

NewEarth, a novel approach to Game Based Learning

In this chapter we address the pedagogical requirements that motivated some choices and how they are reflected in the overall design of the system. In fact, the design has been guided by the pedagogical issues and 21st Century skills discussed so far. Furthermore we have investigated narratology in digital games and Gardner’s theory of multiple intelligence. The final outcomes of this chapter are clear choices that guided the design and prototyping of NewEarth.

3.1 Pedagogical aims

The pedagogical aims of NewEarth are to encourage the acquisition of meta competences and 21st century skills, beside learning STEM in a playful way. We believe that the game should be used as a supporting teaching tool, like a book, and for that reason it should be played in class. The idea is that teachers use the SG together with their students as integral part of the lesson. As a consequence, our SG has been not conceived to be played una tantum and should not be considered as a simulation tool. These requirements are, for us, a natural consequence of the complexity to tackle pedagogical issues described in Chapter 1.

Following that approach, it has been natural to first set a plot, offering an immersive environment, in which the SG can evolve and make the student feel integral part of the game’s mission.

In “Morphology of the Folktale”, Vladimir Propp [83] analyses a selection of hundreds of Russian fairy tales concluding that the form of narrative frames could be summarised in 31 narrative functions. Propp is considered the founder of narratology.
3.1 Pedagogical aims

and his theory of 31 functions have been often used to model plots in digital games, as explained by Sjöström\(^1\). Propp theorises that tales usually start with a description of the situation that is very pleasant and stable, then a disaster occurs that upsets the protagonists’ life. Heroes are then in charge of finding a solution and bringing back the situation to normality.

Inspired by Propp, the first basic choices that we have made in our tale are: people are living in peace (stable and pleasant situation), then a meteorite will impact the Earth (a disaster occurs), scientists should find a new planet suitable for human life (students becomes heroes with a specific mission to accomplish).

Since the beginning of our research, we have chosen to set the plot choosing one of the hottest topic in astronomy: exo-planets and the search of extra terrestrial life. Since the SG wants to address STEM subjects, astronomy and astro-biology perfectly support interdisciplinary and cross topics discussion. Exo-planets are planet orbiting a star other then the Sun, with characteristic similar to the Earth. There is a lot of research in this field that passionsates both adults and young people. Recently (22/02/2017) NASA’s Spitzer Space Telescope has revealed the first known system of seven Earth-size planets around a single star. Three of these planets are firmly located in the habitable zone, the area around the parent star where a rocky planet is most likely to have liquid water. The discovery sets a new record for greatest number of habitable-zone planets found around a single star outside our Solar System. All of these seven planets could have liquid water - key to life as we know it - under the right atmospheric conditions, but the chances are highest with the three in the habitable zone\(^2\).

“This discovery could be a significant piece in the puzzle of finding habitable environments, places that are conducive to life” said Thomas Zurbuchen, associate administrator of the agency’s Science Mission Directorate in Washington. “Answering the question ‘are we alone’ is a top science priority and finding so many planets like these for the first time in the habitable zone is a remarkable step forward toward that goal.”

Figure 3.1. The possible surface of TRAPPIST-1f, one of the newly discovered planets in the TRAPPIST-1 system. Credits NASA/JPL-Caltech.

\(^1\)http://uu.diva-portal.org/smash/get/diva2:629810/FULLTEXT01.pdf, accessed March 01, 2017  
\(^2\)NASA paper, accessed February 23 2017
This incredibly interesting discovery came up just one year after we finalised the SG’s plot: previously outcomes in the exo-planets research field already inspired our choice.

**NewEarth plot**

A meteorite is approaching the Earth and very soon there will be an impact provoking earthquakes and tsunamis that will devastate continents. Students are part of a space mission and each one of them has a specific role with assigned tasks. The team’s goal is to identify, among three, an *exo-planet* that is suitable for human life. The single objectives are related to the specific role. For example, physicists have to calculate the gravity on the three planets and compare it to the Earth’s gravity. Chemists have to derive the planet’s atmosphere composition and so on. Once the student/player has accomplished his/her tasks, the team should compare the obtained individual results and discuss together to infer which is the best planet among the three available ones.

Another important decision that we have made, from a pedagogical perspective, is to implement *NewEarth* as a **Role-Playing Game**.

**Role playing game**

RPGs exist in many forms, from virtual role-playing, to tabletop, to live action. While each type of role-playing offers a unique experience, these games provide a compelling escape from the mundane reality, attracting millions of players worldwide. Unlike the passive experience of watching a film or reading a book, these games encourage players to actively take part in the adventure, sometimes even developing their own stories and characters. RPGs also offer a safe, relatively consequence-free space where players can develop certain aspects of themselves. Through role-playing, players learn how to inhabit the head space of someone other than their primary ego identity, offering them the chance to develop a stronger sense of empathy. The shared, formative experience of RPGs provides a ritual atmosphere for players to enact compelling stories or perform unusual, extraordinary deeds. In this way, RPGs help encourage a sense of community, by teaching individuals to function as a group. Experiences transpiring in RPGs allow players to develop a deeper understanding about themselves and one another during the course of the adventure, as explained by Bowman in [10].

Another important factor that leads to the decision of implementing the SG as a RPG is the collaborative aspect of learning, that is one of the 21st century skills. Borich [8] asked, “What good are critical thinking, reasoning, and problem-solving skills if your learners cannot apply them in interaction with others?” Cooperative learning activities instil important behaviours into learners that prepare them to reason and perform in an adult world, attitudes and values of learners are formed through social interaction. Most of our attitudes and values are formed by discussing what we know or think with others and our attitudes and values are among the most important outcomes of schooling. They provide the framework for guiding our actions outside the classroom. Cooperative learning is important in helping learners
acquire from the curriculum the basic cooperative attitudes and values they need to think independently inside and outside of the classroom. We believe that RPG can strongly support collaborative learning, exactly for their intrinsic nature to develop a deeper understanding of other players.

In support of RPG, we have also investigated a very important study, Dickey in [28], about Massive Multiple Online RPGs (MMORPGs). That study found out that the structure in MMORPGs might inspire the design of interactive learning and game-based learning environments by looking at the elements that support intrinsic motivation. Specifically, they found that the intrinsic motivation can be increased by both character development, role-playing and the small quest design in the narrative environment. In Table 3.1 we report what suggested by Dickey about intrinsic motivation and MMORPGs compared to what NewEarth offers.

Table 3.1. Intrinsic motivation in MMORPGs design vs NewEarth

<table>
<thead>
<tr>
<th>Intrinsic motivation</th>
<th>MMORPGs design</th>
<th>NewEarth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Choice</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character design</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Traits, skills, etc.</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Narrative environment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>small quest</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narrative environment</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Quest selection</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Order of completion</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Strategies employed</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Collaboration</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character design</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Social vs player’s character</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Narrative environment</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Chat and communication tools</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Collaborative quests</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Challenge</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character design &amp; narrative env.</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Quests, level of skills, attributes</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Achievement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character design &amp; narrative env.</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Marked progress indications</td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

To further investigate the process of learning and related theories, we have explored Howard Gardner theory of multiple intelligences.

Theory of multiple intelligence and entry points

The theory of multiple intelligences differentiates intelligence into specific 'modalities', rather than seeing intelligence as dominated by a single general ability. Howard Gardner proposed this model in his 1983 book *Frames of Mind: The Theory of Multiple Intelligences* [19]. According to Gardner, an intelligence must fulfil eight criteria: potential for brain isolation by brain damage, place in evolutionary history, presence of core operations, susceptibility to encoding (symbolic expression), a
distinct developmental progression, the existence of savants, prodigies and other exceptional people, support from experimental psychology and psychometric findings. Gardner chose eight abilities that he held to meet these criteria: musical-rhythmic, visual-spatial, verbal-linguistic, logical-mathematical, bodily-kinesthetic, interpersonal, intrapersonal, and naturalistic. He later suggested that existential and moral intelligence may also be worthy of inclusion. Although the distinction between intelligences has been set out in great detail, Gardner opposes the idea of labelling learners to a specific intelligence. Gardner maintains that his theory of multiple intelligences should empower learners, not restrict them to one modality of learning: an intelligence is “a biopsychological potential to process information that can be activated in a cultural setting to solve problems or create products that are of value in a culture.” Together with his theory of multiple intelligence, Gardner formulated his theory about entry points [38], that specifically deals with the process of learning:

“My own belief is that any rich, nourishing topic - any concept worth teaching - can be approached in at least five different ways that roughly speaking map onto the multiple intelligence. We might think of the topic as a room with at least five doors or entry points into it. Students vary as to which entry point is most appropriate for them and which routes are most comfortable to follow once they have gained initial access to the room. Awareness of these entry points can help the teacher introduce new materials in ways in which they can be easily grasped by a range of students; then, as students explore other entry points, they have the chance to develop those multiple perspectives that are the best antidote to stereotypical thinking.”
Table 3.2. Gardner’s theory of entry points vs NewEarth

<table>
<thead>
<tr>
<th>Howard Gardner Entry Points</th>
<th>NewEarth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narration entry point (read or tell a story)</td>
<td>The game has a plot (tell a story) that evolves during the game</td>
</tr>
<tr>
<td>Logical-quantitative entry point (provide data, use deductive reasoning, examine numbers, narrative plot structure, cause and effect relationship)</td>
<td>During the game, students should solve problems and specific assigned tasks.</td>
</tr>
<tr>
<td>Philosophical entry point (big questions about reasoning and the way of reasoning)</td>
<td>Students should consider pros and cons of every possible solution. They should discuss all together and understand the implications of their choices.</td>
</tr>
<tr>
<td>Aesthetic entry point (emphasise sensory, activate aesthetic sensitivities)</td>
<td>The information/social space is explicitly represented as a 3D immersive world.</td>
</tr>
<tr>
<td>Experimental entry point (hands-on-approach, dealing directly with materials, simulation, personal explanations)</td>
<td>The game requires that students take actively part to the story by solving problems and finding solutions. The team discussion is also a must.</td>
</tr>
</tbody>
</table>

Table 3.2 reports the five entry points theorised by Gardner (first column) and how these entry points are satisfied by NewEarth (second column). We believe that each entry point theorised by Gardner can be perfectly contextualised in a RPG, where each student can find the entry point(s) that best fits for him/her.

All the assumptions made so far make NewEarth fully compliant with what reported in Figure 1.1: the pedagogical issues that entered the game design, automatically lead to the development of 21st century skills, as explained hereafter.

21st century skills

To address the problem solving skill, players in the game have to solve complex problems (like habitability of the planets) broken into simplified problems (like calculating single key parameters, e.g. gravity).

Critical thinking is reinforced by the nature of the problems in the games: for example, planets’ characteristics should be compared with Earth’s data. In addition, a complete vision of the overall solution, starts coming out after discussion and critical thinking: ”Why a planet is better then another?”, ”The presence of toxic gases makes the atmosphere compliant with human life?”, ”What happens in a different atmospheric pressure?”, ”Is there liquid water on the planets?”. All these questions are not directly posed to students but are revealed to students as they start seeking for a solution to the initial quest ”Which is the best exo-planet suitable for human life?”. The initial quest is the catalyst for further questioning and for an inquiry-based approach. Inquiry-based learning starts by posing questions, problems or scenarios, rather than simply presenting established facts or portraying a smooth
3.1 Pedagogical aims

path to knowledge. Inquirers will identify and research issues and questions to develop their knowledge or solutions. The inquiry-based instruction is very closely related to the development and practice of thinking skills.

*Collaboration* and *communication* are natural outcomes of RPGs and are specifically addressed by *NewEarth*. In fact, players beside collaborating to reach common objectives, can share their discoveries with teachers and other players during the game. The latter, inspired by the Knowledge Sharing paradigm, has been designed ad hoc for *NewEarth*, as explained in Section 4.1.1.

**Interdisciplinary**

Another important pedagogical aspect, that found easy application in RPG, is bound to *interdisciplinary* issues. In the last few decades, many reform initiatives have shaped teaching and learning in Science. These reform efforts include a shift from teaching students to remember and execute isolated facts and skills, to having students experience learning as scientists, engineers and mathematicians do. Pedagogues argue that students should engage in learning that allows them to explore, inquire, solve problems, and think critically. To this end, reform efforts within each of the STEM disciplines have focused on such strategies as inquiry learning [100], project-based learning, constructivist learning [67], problem-based learning [40] and the integration of technology across all STEM disciplines. Although these efforts have fostered improved learning outcomes within each of the STEM disciplines [18], many researchers argue that in order for students to be fully prepared for careers in the new millennium, they must be capable of thinking across disciplinary boundaries, Berry et al. in [6]. Frykholm and Glasson suggest that schools must begin to veer away from treating each STEM discipline as a silo and embrace an approach that blurs the boundaries of these disciplines. They argue that students who engage in rich cross-disciplinary experiences will have a deeper conceptual understanding of science and mathematics content [36], which will improve their achievement in each of the disciplines. Further, interdisciplinary learning can foster an understanding of STEM concepts in their application to real world problems, which are interdisciplinary by nature. In traditional school settings, the compartmentalization of scientific knowledge creates boundaries so rigid that they often serve as barriers to any efforts to develop integrative science and mathematics programs, Nikitina & Mansilla in [75].

The above reasoning leads us to the definition of four roles.

**NewEarth** roles and tasks

*NewEarth* embraces interdisciplinary and integrated science concept. Table 3.3 reports the description and final task for each role, as they appear in the game.
### 3.1 Pedagogical aims

#### Table 3.3. Players’ roles in NewEarth

<table>
<thead>
<tr>
<th>Role</th>
<th>Description</th>
<th>Overall Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicist</td>
<td>You study a wide range of phenomena, spanning all length scales: from sub-atomic particles of which all ordinary matter is made, to molecular length scales of chemical and biological interest, to cosmological length scales encompassing the Universe as a whole.</td>
<td>Which is the value of gravity on the three exoplanets? Make a comparison with terrestrial gravity, discuss with other team members, evaluate if the value you will find is suitable for human life!</td>
</tr>
<tr>
<td>Geologist</td>
<td>You work to understand the history and future of our planet. The better you understand Earth’s history the better you can foresee how events and processes of the past might influence the future.</td>
<td>Axial tilt and magnetic field are important planets’ characteristic for human life? Water cycle is the same in different atmospheric condition?</td>
</tr>
<tr>
<td>Chemist</td>
<td>You master composition of matter and its properties, you deal with molecules and atoms, you carefully measure substance proportions, reaction rates, and other chemical properties.</td>
<td>Can humans breath a different atmosphere? What happens if we find toxic gases like methane or ammonia? How much water do we need to survive in a different habitat?</td>
</tr>
<tr>
<td>Astro-biologist</td>
<td>Your interest is the study of the origin, evolution, distribution, and the future of life in the universe: extra-terrestrial life and life on Earth.</td>
<td>Which are the atmospheric temperature and pressure on the three planets? Check human survival conditions!</td>
</tr>
</tbody>
</table>

Each role has three tasks to accomplish. A detailed description of all the tasks is reported in Appendix B. Another important aspect that we have taken into account in the design phase is to provide an interface for both students and teachers.

**Interfaces for both students and teachers**

In fact, while the majority of SGs offers an interface for students (the serious game itself), they do not offer a dedicated interface for teachers. We believe that teachers should not be excluded or just assist the players. They should be active actors of the game, conserving their leading role of coordinators. For that reason we decided that the teachers’ interface should be integral part of the SG.

In the next section, we analyse some very basic principles of game design and how we used them in the design phase of NewEarth.
3.2 Serious Game Design

There are many ways to break down and classify the many elements that form a game. We found very useful the classification made by Schell in his book: "The Art of Game Design: a book of lenses" [94]. To analyse a game, Schell uses the elemental tetrad, reported in Figure 3.2.

![Game design elemental tetrad](image)

Let’s look briefly at each of the four elements and how they relate to the others.

**Mechanics:** These are the procedures and rules of the game. Mechanics describe the goal of the game, how players can and cannot try to achieve it, and what happens when they try. When the designer chooses a set of mechanics as crucial to the gameplay, it is crucial to choose technology that can support them, aesthetics that emphasise them clearly to players, and a story that allows the game mechanics to make sense to the players.

**Story:** This is the sequence of events that unfolds in the game. It may be linear and prescripted, or it may be branching and emergent. When there is a story to tell through the game, mechanics should both strengthen that story and let that story emerge.

**Aesthetics:** This is how the game looks, sounds, smells, tastes, and feels. Aesthetics are an incredibly important aspect of game design since they have the most direct relationship to a player’s experience. When the game has a certain look, or tone, that players should experience and become immersed in, it is very important to choose a technology that will not only allow the aesthetics to come through but amplify and reinforce them. Mechanics should make players feel like they are in the world that the aesthetics have defined, and the story should evolve with a set of events that let the aesthetics emerge at the right pace and have the most impact.

**Technology:** The technology enables the game to do certain things and prohibits it from doing other things. The technology is essentially the medium in which the aesthetics take place, in which the mechanics will occur, and through which the
story will be told.

In Table 3.4 we give a very short description of the four elemental tetrad that we defined in *NewEarth*.

<table>
<thead>
<tr>
<th><strong>Aesthetic</strong></th>
<th>The aesthetic is bound to the environment. We have a spaceship laboratory plus three simulation rooms, one for each planet to be studied.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Story</strong></td>
<td>Players have to study three exoplanets and decide which is the best to inhabit.</td>
</tr>
<tr>
<td><strong>Mechanics</strong></td>
<td>Players should solve an initial quest. In order to do it, they should explore the environments, collect objects, solve quizzes and tasks. For a detailed description see also Table 3.5.</td>
</tr>
<tr>
<td><strong>Technology</strong></td>
<td>Game engine, web services, AI algorithms.</td>
</tr>
</tbody>
</table>

We can consider both games and courses as a system of problem-solving activities. Problem-solving is a process of activities aimed at transitioning from an initial state to a desired goal state, overcoming obstacles, and often developing knowledge, Smith et al. in [98]. Some types of games explicitly leverage the problem-solving nature of gaming in order to articulate the player experience. This is the case of quest-based games, in which quest systems are used as mechanics to organise game play activities and contextualise them in a coherent narrative frame, defining narrative and activity progression within the game, as explained by Howard in [46]. Quest-based games have a primary aim that players must fulfil through achieving concrete objectives associated with quests. Thus, quest-based games are deliberately designed as systems of problem-solving activities aimed at achieving objectives that need to be successfully addressed in order to progress and eventually win. Gameplay mechanics defines the interrelationship and interoperability of the game system elements, the provision of contextual information to the player, and the dependencies between gameplay activities, Salen & Zimmerman [90]. We have taken the patterns proposed by Fabricatore & Lopez in [33], as the base for our gameplay development, and extended them with two more patterns: the quiz structure and the team challenge-based reward, as explained in Table 3.5.
<table>
<thead>
<tr>
<th>Gameplay Patterns</th>
<th>Explanation</th>
<th>NewEarth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quest structure</td>
<td>The game follows a quest structure. Individual quests solutions lead to the solution of team’s quests.</td>
<td>Quests are of two types:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. Simple problems that should be solved individually (tasks)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Complex problems that require the work of the entire team, possibly coordinated by teachers.</td>
</tr>
<tr>
<td>Quiz structure</td>
<td>Quizzes are used to introduce players to assigned task.</td>
<td>Players should answer quizzes to gain access to the planets’ simulation rooms</td>
</tr>
<tr>
<td>Orientation</td>
<td>We provide tools that are intended as explicit helps for players. Helps usage increase players’ score</td>
<td>Players can ask questions to the IPA or have access to insights. Asking questions to IPA increases players’ score.</td>
</tr>
<tr>
<td>Individual challenge-based reward</td>
<td>In each game session, players can gain points and increase their personal score.</td>
<td>See Table 4.1 for point assignment</td>
</tr>
<tr>
<td>Team challenge-based reward</td>
<td>Teams’ members should discuss and provide a common solution to the overall objective.</td>
<td>Individuals players’ score in the same team are summed, the result is the team’s score</td>
</tr>
</tbody>
</table>

*NewEarth* game mechanics can be summarised as follows. In order to achieve the overall objective, players should explore the 3D environment, which is formed by a spaceship and three simulation rooms, one for each planet. Players should collect several objects around the ship and gain access - answering correctly to quizzes - to the planets’ simulation rooms. In the planets’ simulation rooms player should solve the assigned task using interactive panels. They can seek for help in any moment, namely asking questions to the IPA or simply consulting some ad-hoc prepared materials (like video, articles, insight, simulation, etc.) external to the game but integral part of the VLE.

As it will be clear in Chapter 4, the VLE coincides with the SG and spans trough different architecture layers. As explained in Chapter 2, Section 2.1, VLEs can be roughly divided into three main areas following the categorization proposed by Dalgarno & Lee in [23], that are:

1. 3D simulation and microworlds
2. 3D environments as interfaces to learning resources
3. 3D multi-user

Due to the complexity of our architecture and to the different components that form it, we can state that our VLE is cross-domains, with respect to the above classification. In fact, our VLE, is formed by a 3D virtual world that students/players can explore and interact with. Beside that, the 3D environment offers the possibility to interfaces to learning resources, that are integral part of the 3D environment: players should solve problems via interactive panels that are part of the 3D virtual world. Moreover, the VLE offers access to resources external to the 3D world: in the game there is the possibility to access insights prepared ad-hoc for the game. Finally, the 3D environment is multi-users but not multi-players, which means that more users can play simultaneously, but they cannot virtually meet in the 3D environment. To conclude the game, players have to send individual solutions via game panels and they have to produce a written report (one per team) that they can upload to the VLE.

In Section 2.1, we reported the rules designed by Dillenbourg to define a VLE. The VLE that we defined satisfies these rules, as reported in Table 3.6.

<table>
<thead>
<tr>
<th>VLE rules</th>
<th>NewEarth</th>
</tr>
</thead>
<tbody>
<tr>
<td>A virtual learning environment is a designed information space</td>
<td>Information is accessible in multiple ways (asking help to IPA, consulting insights, interacting with game panels). Content, can be added by teachers and experts.</td>
</tr>
<tr>
<td>A virtual learning environment is a social space</td>
<td>Players can interact via chat and share ideas with teacher an other players</td>
</tr>
<tr>
<td>The virtual space is explicitly represented</td>
<td>The student virtual space is a 3D virtual environment, namely a spaceship. The teacher virtual space is a virtual board in which teachers can see players information in real time</td>
</tr>
<tr>
<td>Students are not only active, but also actors</td>
<td>Students have roles, individual and team tasks</td>
</tr>
<tr>
<td>Virtual learning environments are not restricted to distance education</td>
<td>NewEarth is conceived to be played in class as a supporting teaching tool</td>
</tr>
<tr>
<td>Virtual learning environments integrate multiple tools</td>
<td>NewEarth supports learning, gaming, collaboration, information, etc.</td>
</tr>
<tr>
<td>Most virtual environments overlap with physical environments</td>
<td>NewEarth offers several experiments that very often are presented in real STEM labs</td>
</tr>
</tbody>
</table>
3.3 Research open questions

In recent years there has been an upsurge of interest in the use of digital games for learning and behaviour change with an increasing number of conferences, journals, projects and societies devoted to the topic. Much of this interest has been speculative, discussing the potential of games to provide new methods for supporting learning. To address concerns about the lack of empirical evidence about the effectiveness of games, Boyle et al. (see [11]) carried out a literature review that aimed to identify research evidence about the positive impacts of games. The review confirmed that research on games was very diverse with respect to the focus of the studies, the outcomes reported, underlying theoretical models and methodological approaches. In that study, authors advocated a narrative review to capture the diversity of the selected studies and they developed a multi-component analysis of games and their outcomes which provided a useful framework for organising the research along key variables, making links between the different ways of categorizing games and outcomes, and identifying emerging trends, hot topics and gaps in the literature. Authors found more than 500 papers reporting empirical evidence of the positive outcomes of playing games. However, there is still place for improvement, in fact from the study emerged the following open questions.

**Open issue:** Pedestrian use of Serious Games. In [11], Boyle et al. state that games for learning are still most frequently used to support knowledge acquisition and do not exploit the potential of SGs to teach 21st century skills and behaviour change.

The majority of SGs are frequently championed as novel, engaging and active new method for supporting 21st century skills and behaviour change. However very often SG are used like other standard teaching tool, as a book for example, and do not exploits all the potentiality that the SG offer. We put a lot of effort, during the design and implementation phase, to address the above mentioned 21st century skills. In Chapter 6 we tried to specifically assess the novelty of our SG in this direction.

**Open issue:** Retention and long-term learning. In [66], De Marcos et al. suggest that gamification and Serious Games foster shallow learning, but ultimately the question of retention and how SGs support or undermine long-term learning still remains open and further enquiry is needed.

In our design and development cycle, we developed a stand alone application (namely and Android app) to test the functionality of our VT, before integrating...
it into the SG. We tested the retention of the acquired knowledge and results are reported in Chapter 6, Section 6.2.
Chapter 4

**NewEarth architecture & design**

In this chapter we explain the overall system architecture (3-tiers) and the main components that form it. In particular we have implemented two interfaces for teachers, that should actively participate to gaming sessions, and one for students, the 3D virtual game. Beside the presentation layer and the afore mentioned application, we introduce the logic and data layers as well. In the last section of the chapter, the IPA design is described in details.

Recapping from the previous chapter, in our aims, the game should satisfy the following pedagogical requirements:

- **It should be used as a supporting teaching tool**
- **It should be a role playing game**
- **It should address some of the requirements for the 21st century skills**
- **It should foster interdisciplinary and collaboration**
- **It should offer an interface both for students and teachers**

As we anticipated in Section 1.4, we have implemented a VLE as a SG in which an IPA guides the player trough the game. Thus, from an engineering point of view, the game:

- **should satisfy the characteristics of VLE**
- **should integrate an IPA**

In Figure 4.1 we have represented the basic concepts and technologies that lead to the architecture design: from one side there are the pedagogical requirements and
from the other side there are the technical requirements, which combined together lead us to the architecture design.

Figure 4.1. Architecture design procedure

The three-tier overall architecture that we have designed is depicted in Figure 4.2. A three-tier architecture is a client-server software architecture in which the user interface (presentation), functional process logic (business rules), data storage and data access are developed and maintained as independent modules, most often on separate platforms. Apart from the usual advantages of modular software with well-defined interfaces, the three-tier architecture is intended to allow any of the three tiers to be upgraded or replaced independently of the others, in response to changes in requirements or technology.

Figure 4.2. Overall architecture

The first one is the presentation tier, which consists of three users’ applications, one for students (3DVG), which runs on mobile devices (tablets) and two for teachers (PA & PVB) that run, respectively, on the web and on a multimedia whiteboard.
run, on a multimedia whiteboard and on the web.

- **Teacher workspace:**
  1. Professor App (PA)
  2. Professor Virtual Board (PVB)

- **Student workspace**
  1. 3D virtual game (3DVG)

The logic tier is pulled out from the presentation tier and controls the application’s functionalities by performing detailed processing. It offers the services that permit us to administer players and teachers (like register, login), to store information about players (using a dedicated database), to manage teams, to manage game’s sessions, etc.

The data tier includes the data persistence mechanisms (database servers, file shares, etc.) and the data access layer that encapsulates the persistence mechanisms and exposes the data. The data access layer provides APIs to the application tier that exposes methods of managing the stored data without exposing or creating dependencies on the data storage mechanisms. Avoiding dependencies on the storage mechanisms allows for updates or changes without the application tier clients being affected by or even aware of the change. The Data Tier hosts a database for storing all information related to users (teachers and students) and to the IPA knowledge.

Figure 4.3 represents the overall architecture in which we adopted another client-server perspective: the logic and data tiers are on the server side, while the presentation tier is on the client side. The client side hosts the teacher and student workspace.
4.1 VLE

4.1.1 Presentation tier: teachers’ workspace

The teacher workspace has been designed for teachers in order to actively participate to games’ session and, at the same time, to analyse students/players results both while playing and in other moments of their working activities. Teachers’ workspace consists of two different applications designed and developed by us, plus an off the shelf application named Opedia®. The latter was used to share multimedia contents among teachers and students, as pictured in Figure 4.4:

- Professor App
• *Professor virtual Board*

The PA has been conceived to start/stop game’s sessions (from now on just session) and to view collected data about students’ activities after sessions. It can be accessed any time anywhere. The PVB should be used in class; it is explicitly conceived for handling sessions, it is interactive and it has been designed to run on a multimedia whiteboard.

![Figure 4.4. Teacher workspace](image)

As we will see in Chapter 5, the development of the two applications has been totally different. In fact, PA is very essential and spartan while PVB offers a nice game interface, in line with the student’s one.

**Professor App.**

It permits to teachers to handle schools, classes and game sessions. In addition it permits to visualise results of previous and ongoing game sessions

As already mentioned PA is deliberately simple. Teachers could access it from a PC but also from their personal mobile devices, like phone or tablet. They can manage the classes where they teach and where they want to use the game. More in details, teachers can:

- Register to the PA.
- Login to the PA if already registered

Once they are logged in they can:

- Select the academic year and choose the schools where they teach
- Add/Select a class where they teach

Once classes are already registered they can manage the session:

- Start/stop game session for each class.
- See students’ results in ongoing or previous sessions
Once a teacher starts a session, a unique code (PIN) is generated. The game session lasts four hours. This PIN should be used to start the PVB component in class. This mechanism prevents two teachers that teach in the same class from having conflict during game sessions and permits them to have several contemporary game sessions, for the same class.

Detailed use cases are provided in the logic tier description (see Figures 4.5 and 4.6).

### Professor Virtual Board

It runs on an interactive whiteboard. The aims of this component are both to give an instrument for teachers to actively take part in ongoing sessions and for students, who can check what other team’s members and competitor teams are doing.

More in details teachers can:

- Launch the PVB.
- Insert the PIN already generated using the PA.

Once the app is started they can:

- Add new teams

Once new teams are created, students should register from their devices and connect to a specific team (see next section on 3DVG). After that, teams are automatically updated with the registered players. That operation needs to be done just the first time that teacher launches the PVB for that particular class. In the next sessions, teams are already formed and teacher can, as depicted in Figure 4.8:

- Add new teams if necessary
- Access players’ information (like role, assigned task, score)
- Check players’ position on the map
- Validate or not players’ note in the Knowledge Forum.

The last point concerns the possibility, for the players, to add notes and share them with the team. Notes are simple text messages containing information about a players’ discovery in the game. Information can be of any type, like "I found water on planet Maya", "Temperature on Melissa is very high", etc. Once notes are sent, they appear in the PVB interface. Teachers have the option whether or not to validate the note. If the note is validated players gain points (see section on score) and the notes is stored in the team’s Knowledge Forum. The idea to share discoveries and pieces of information was borrowed by the knowledge sharing area of research and in particular was inspired by the Knowledge Forum product.

---

1 Knowledge Forum provides students and teachers with a unique collaborative space in which to share ideas and data, organise course materials, analyse research results, discuss texts, and cite reference material. http://www.knowledgeforum.com/index.htm
4.1 VLE

**Opedia®**

It is an off the shelf product that has been integrated into the VLE. It permits to share files, video, links, etc. among teachers and students.\(^2\)

4.1.2 **Presentation tier: Students’ workspace**

The student workspace consists of one application:

- 3D Virtual Game

**3D Virtual Game**

The 3DVG is the core part of the students’ workplace. Students become players when they first register to the game, hook to an already existing team for their class, and choose a role among the four described in Table 3.3.

Each player belongs to a team. Once the teacher starts a new session, players can start playing. In particular they can:

- login to the 3DVG
- explore the 3DVG
- recover 3D objects they need to solve the assigned task
- gain access to the simulation’s rooms answering quizzes
- enter the simulation’s rooms and find the interactive panels
- solve assigned tasks using interactive panels
- ask help in natural language to the IPA
- access Opedia and insights
- add notes
- chat with other team’s members

All the actions listed above can be performed by players in the sequence they like, following their own strategies. Game mechanics just force some sequence of actions. In fact, it is not possible to access the simulation rooms if players have not answered correctly quizzes. It is possible to enter the simulation rooms without having collected all the objects, so the player can explore the 3D environment. However, players cannot access the interactive panels to solve the assigned task if they have not collected the necessary objects. Beside that, the player has to guess which are the right objects for the assigned experiments, choosing from the inventory.

---

\(^2\)We will not discuss Opedia® in this thesis, we just stress that this off the shelf platform has been integrated into our VLE thanks to an agreement that we have made with its producer.
This is part of the critical thinking skill: player have to think about the experiment he/she are going to do and how to interact with it. Most of the actions in the game permits to gain points.

The score mechanism is an important part of the 3DVG because score deals with students assessment. After a debate with some pedagogy experts we decided to have just positive score and not negative. So, if players do something wrong, like for example sending a note containing jokes, the score remains unchanged instead of decreasing. That occurs because the negative score can be experienced by players as a discouraging experience in the overall game.

In the 3DVG, players can gain points following the schema proposed in Table 4.1. Those points have been decided following the assessment used in school. Quizzes are simpler then tasks and so they give fewer points. Adding notes and asking questions have an important weight in the score because we want students to use them. Badges are inspired by commercial games and thus are not related to assessment but just to reward students for completing various activities, which are one of the key lever in game mechanics. During the game, players are notified about their successful actions but they do not know the amount of points they have earned. Points can be seen just asking teacher to check the score in the PVB, which encourage an additional interaction between students and the teacher. Players, beside points, gain badges that can be accessed in the trophy room, that is one of the 3D virtual environments. That decision was made to encourage a contact between players and the teacher during a game session. On the PVB, it is possible to check also the team score.

Table 4.1. NE scores and badges

<table>
<thead>
<tr>
<th>Action</th>
<th>Score</th>
<th>Badge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer quiz correctly</td>
<td>1st attempt, 5 points 2nd attempt, 3 points 3rd attempt, 1 point</td>
<td>1 badge</td>
</tr>
<tr>
<td>Solve individual task</td>
<td>1st attempt, 10 points 2nd attempt, 7 points 3rd attempt, 3 point</td>
<td>1 badge</td>
</tr>
<tr>
<td>Find all objects</td>
<td>5 points</td>
<td>1 badge</td>
</tr>
<tr>
<td>Add note</td>
<td>2 points per note, maximum possible notes five</td>
<td>1 badge</td>
</tr>
<tr>
<td>Ask question</td>
<td>3 points per question, maximum possible questions five</td>
<td>1 badge</td>
</tr>
<tr>
<td>Chat</td>
<td>2 points, just first access</td>
<td>1 badge</td>
</tr>
<tr>
<td>Access Opedia</td>
<td>2 points, just first access</td>
<td>1 badge</td>
</tr>
</tbody>
</table>

4.1.3 Logic Tier

The logic tier offers all the services that permit to administer players and teachers (like register, login), to store information about players (using a dedicated database),
to manage teams, sessions, etc. The logic tier communicates via network with the presentation tier and the data tier.

Several use cases are reported hereafter, according to the following sequence of actions:

1. Students and teacher register in the environment
2. Students and teachers log into the system
3. The first time students are logged in, the teacher explains the game, roles, tasks, etc. Each student is assigned a role and teams are formed. From now on students become players.
4. Players play the game.
5. Players can use several functionalities offered by the 3DVG interface.

The following use cases describe most of the functionalities offered by the logic tier. We have reported just the more significant ones.

<table>
<thead>
<tr>
<th>Use Case 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Users register. Users can be teachers or students.</td>
</tr>
</tbody>
</table>

Students and teachers register via the user interfaces (that are respectively, 3DVG and PA) on a mobile device; they should provide mandatory data like: name, surname, email, school, class, password, role. Information is registered in the DB and a unique identifier, ID, is assigned to each user. It is important to stress that a student can register just after the teacher has created teams for the class. In fact, at registration time students can select a team and hook to it.

![Diagram of user registration process](image)

**Figure 4.5.** Users register into the system
Use Case 2

Teacher starts/stops game session for a specific class

From PA interface, teachers can start/stop a session. When a session is started a unique PIN number is generated. This PIN number is returned to the teacher and should be used to launch the PVB.

![Diagram](image)

**Figure 4.6.** Teacher starts/stops game session

Use Case 3

Teacher forms teams: choose a name and add team to the specific class.

The teacher launches the PVB from the multimedia interactive whiteboard in class. The teacher has to insert the PIN, previously generated using PA (see Use Case 2). The PIN is unique for the class and lasts for four hours. After that time has passed, the teacher has to generate a new PIN. From the PVB, the teacher adds teams. Teams’ names can be chosen from either the teacher or the students. They are stored in the DB and hooked to the specific class. This mechanism is very important because the game can be played in different lessons. Once a teacher logs in again, for that specific class, with a different PIN, he sees the already created list of teams. Then, he can simply follow the game or decide to add new teams if necessary.
Use Case 4

*Players log in and play the game.*

As already specified, once students have registered for the first time to the 3DVG (see use case 1) they become players. Students should simply provide their data (email, password) and enter the game environment: the game session is started. From now on, players can simply play the game, using all the features offered by the 3DVG.

Use Case 5

*Teacher launches the PVB and follows the game session’s progresses*

Once a teacher has already created teams and players are already registered, he/she can follow game progression on the PVB. In particular, he/she can follow players position on 2D map of the 3D environment, score progression and validate notes.
Use Case 6

Add notes to Knowledge Forum

The Knowledge Forum gives to students the possibility to add notes about any learning aspect of the game. The notes appear on the PVB and are visible to everyone, the teacher can validate them or not.
Use Case 7

Use Case 7. Player asks a question to the IPA

When players ask a question to their IPA, the AI algorithm checks the database and performs a best match (as extensively explained in the paragraph about DMS). Answers can be of type normal or no-match. When type normal is returned, the IPA provides an answer back to the player, if no-match type is returned, the IPA asks to repeat the question or interact with an interlocutory sentence.
4.1 VLE

4.1.4 Data Tier

The data tier consists in a complex database where all data, that are related to the different actors of the system, are stored. In particular, the involved actors in our system are:

- Users. They can be teachers, students and players.
- NPC (Non-player Character\(^2\)).

We may have an unlimited number of users and an unlimited number of NPCs. In our case, we have just four NPCs, one for each role. In fact, the association between a player and a NPC is done by role. The database plays a key role also for the communications among the different components of the system: data are stored in the database from one component and read from another one.

To design the database we used an entity-relationship model (ER model), which describes inter-related things of interest in a specific domain of knowledge. An ER model is composed of entity types (which classify the things of interest) and specifies relationships that can exist between instances of those entity types. The whole purpose of ER modelling is to create an accurate reflection of the real world in a database. The ER model does not actually give a database description but it gives an intermediate step from which it is easy to define a database.

\(^2\)In a game, any character not controlled by a player is referred to non-player character (NPC). In video games, this usually means a character controlled by the computer through artificial intelligence. From now on we refer to the IPA as NPC, in relation to the DB functionalities.
The relation schema that we designed is reported in Figure 4.11

![Database ER diagram]

Figure 4.11. Database ER diagram

In the left side of the above figure we have the generic entity *user* with all its attributes. Users can be *professor* or *player*. Player is an instance of *student*. A student pertains to a *class*, is part of a *team* and is associated with a *character* (role). Students can post *note* which is an entity with attributes like topic and text. Professors teach in one or more classes. Classes belong to *schools* and teams belong to class. In the right side of the picture we have the *IPA*. Each IPA has its own attributes like name, subject and description. Left and right side are bound by *character* and *player sentence*; the latter represents the question that players ask to IPA. IPA sentence is related to the answers, questions and suggestions. *Answer* can be of two types, normal or no match. When type is normal, the IPA returns the best answer, when type is no match it means that the IPA did not find any answer. In this case the IPA return a phrase like "Sorry, I didn’t get it. Can you ask me a new question?". That mechanism is fully described in the following Section 4.2.1, Algorithm 1.

The database is populated by users when they register to the system for the first time. Some attributes like score are updated during game’s sessions.

The database stores all information related to the NPC. In particular the NPC’ templates consist of simple sequence of question-answer pairs and eventually suggestions. Here it follows an extract of the physicist’s NPC data that we have used to populate the DB, see Figure 4.12.
The NPCs are bound to the characters, which means that each NPC table is populated with ad hoc information. This information is strictly related to the tasks that players should solve in the game and it has been prepared by field expert.

More details about IPA and NPC can be found in the following section, entirely dedicated to the IPA design description.

4.2 IPA

As already mentioned, the IPA is fully part of the 3DVG, and should guide the player through the game, replying to questions asked in natural language. The IPA that we have designed has two abilities. The first one is to reply to question asked...
in natural language: it can be thus classified as a Dialogue Management System (DMS). The DMS that we have implemented is described in Section 4.2.1. The second ability is to evaluate the game progression and emotions expressed by players during the game. Game progression and emotions are used to provide help when needed, as explained in further details in Section 4.2.2. From now on we refer to this research issue as the ability of the IPA to provide unsolicited hints. As we have explained in Table 2.1 that ability makes our Virtual Tutor pedagogically intelligent: we pass from an Intelligent Virtual Tutor implemented as DMS to an Intelligent Pedagogical Agent capable of pedagogical intervention.

4.2.1 Dialogue Management System

The proposed DMS manages dialogues where the player has the ability to interact, in natural language, with NPC. It is important to allow the interaction through natural language, since this gives the perception of a deeper involvement with respect to traditional communication systems. As a matter of fact, while in state-of-the-art games, the user usually interacts by choosing a sentence among a set of predefined possibilities, our system allows the user to freely express his thoughts in textual form and provides the user with an adequate answer (selected from a database of answers provided by a field expert). Especially in serious games, it is important to stress the feeling of immersion in the game action. While in videogames high engagement levels increase pleasantness and pure entertainment, in the case of serious games, raising the level of immersion allows better transmission of the game message, Cartelli in [15], which is a didactic message.

In our system, the communication is implemented through an NLP algorithm based on an ad hoc text retrieval problem solver and on a Naïve Bayes text classifier with an inner product-based threshold criterion. Our DMS addresses a specific target of human-content interactions. In particular, these kinds of NPCs may be employed in short dialogues where a player could get knowledge clues about a specific topic. Thus, such NPCs would implement a specific serious game mechanism combining user knowledge acquisition and natural, pleasant interaction. In general, SG are not interactive dramas, such as the well-known Facade ³, where dialogue is the key element. In state-of-the-art serious games, the user text/speech interaction is limited.

We have implemented a variation of a text retrieval algorithm. The problem of selecting the right answer from a Knowledge Database (KDB) can be expressed formally similarly to an information retrieval problem, Lease in [59]. This expression refers to a problem where we need to query a collection of documents with the goal of identifying the most suitable document to answer to the user’s query. For each query, the system performs a comparison between the query and the collection’s documents, classifying them according to the relevance of the document with respect to the query. The system will give back to the user the document with the highest relevance value. In our case, the user’s query corresponds to the user’s question, and the documents of the collection are represented by the sentences stored in the KDB.

There are three approaches to the solution of a text retrieval problem.

³http://www.interactivestory.net/
The first is the classical approach, which consists in the evaluation of the relevance of a document in terms of vector similarity between the query and the collection of documents. This model is also referred as the Vector Space Model, Singhal in [96], since both the query and the documents are represented as vectors in a high dimensionality space, in which every component relates to a word of the sentence. By performing the dot product of the two vectors, it is possible to measure how much the two vectors diverge or tend to overlap. This method needs a criterion to assign an appropriate weight to each vector component. It is based on the computation of the two statistical parameters Term Frequency (TF), which measures the relative frequency of the term in the document, and Inverse Document Frequency (IDF), which expresses the importance of the term within the whole collection of documents.

The second approach is a probabilistic one, which relies on the Probability Ranking Principle by Robertson [87], according to which the documents of a collection should be sorted with respect to the probability of being relevant to the query. This can be obtained through a binary classification in two classes relevant and not relevant.

The third approach relies on the Language Modeling Theory, as introduce by Ponte & Croft in [82], which makes use of the hypothesis that every document in the collection derives from a unique Language Model. The computation of the document’s relevance value is based on the probability of obtaining the user’s query as a drawing of a sample randomly generated by the language with which the document has been created.

Basing this research on the work done by Mori et al. [73], we designed a hybrid model where we used both the first and the third approach.

**Text Retrieval Problem: IPA algorithm**

The core of our system is an implementation of Naïve Bayes Text Classification, Manning et al. in [65], which is a probabilistic classification method based on Language Modeling under the hypothesis of words’ conditional independence. This algorithm is an application of the Bayes Theorem, which allows the estimation of the probability that a collection of documents is relevant for a query, given the sequence of words that make up the query itself. Assuming that each document belongs to a class, we can determine the probability that a class is relevant, given the query’s sequence of terms. Moreover it is important to give the user an answer extracted from the KDB only if there is a certain amount of confidence that the chosen answer is appropriate to the query. Otherwise the system has to reject the identified answer and it has to return a sentence telling that the NPC cannot answer the user’s question. We therefore validate the answer chosen by the Naïve Bayes algorithm by applying a threshold criterion, that is a method like the inner product similarity measure, which gives the same numerical result independently of the dataset details. So, after the Naïve Bayes algorithm finds the best match, we assess its appropriateness to the user’s query by applying the inner product between the found result and the query, in order to accept the result only if the dot product isn’t below a certain value.

The algorithm is reported as Algorithm 1. The first task consists in a sentence lemmatization, which transforms all the terms of the sentence in a form suitable
4.2 IPA

for word analysis and computation. We then evaluate the Interrogative/Adjective Pronoun (IAP) of the question, in order to assign a higher classification score to the sentences from the KDB with a matching IAP. Instead of excluding all the sentences in the KB with a different IAP, we just penalise them with a lower classification score; in fact it could happen that an answer might be the best one to respond to the user’s question even if each of the expected questions that were associated with it in the KB had a different IAP. We then proceed with the Naïve Bayes Text Classification, and filter the result by applying the threshold criterion. If the answer passes the test, then it will be returned to the strategic module, otherwise the answer will be replaced with a warning message telling the user that his question didn’t produce any match in the NPC’s KDB.

Algorithm 1 Naïve Bayes Text Classifier Algorithm

Input: input question = Q; NPC stored questions = NPCqs

1: \( Q_{\text{lem}} = \text{lemmitise}(Q) \)
2: \( \text{set score} = \text{evaluate}(Q_{\text{lem}}, NPCqs) \)
3: \( \text{sort by score } Q_{\text{lem}} \)
4: \( \text{for } i = 0; i++; Q_{\text{lem}} \text{ n do} \)
5: \( Pr_i = Pr_i(Q_{\text{lem}}) = \text{Pr\_Bayes\_classifier}(\text{sorted } Q_{\text{lem}}) \)
6: \( \text{if } Pr_i > \text{treshold then} \)
7: \( \text{return ("Normal", NPC\text{Answer})} \)
8: \( \text{else} \)
9: \( \text{return ("No match", \text{random("Interlocutory")})} \)
10: \( \text{end if} \)
11: \( \text{end for} \)

In the next paragraph, we explain how we have introduced users’ emotions analysis, advancing the DMS proposed by [73].

4.2.2 Emotions and schooling

As mentioned, one of the abilities of our VT is to catch emotions in real time. In all the papers reviewed in the state of the art (cfr. Section 2.2.2 and Figure 2.2), the suggested approach to catch users’ emotions is top down rather than bottom up. In fact the majority of the papers in this field starts from the theoretical study of emotions, trying to find the best schema in the psychological field (that could be dimensional rather than categorical) and shift this schema in an engineering framework using sensors to catch the real time emotions. We moved away from that approach in two ways: we did not use sensors and we did not use a psychological schema. In the following paragraphs we introduce the motivations for our choices.

Emotions and affective states are pervasive in all forms of communication, including text based, and increasingly recognised as important to understanding the full meaning that a message conveys. In our research we decided to catch emotions in text. This is because \( NE \) has to be played in class and students must feel confident and immerse in a emotionally safe environment. Sensors could not be welcomed by students and we cannot neglect that we deal with minors: tracking biometric
data could represent an obstacle. Another reasoning that we had made concerns the game evolution. In fact, the game is rather slow, players need reasoning and time to provide solutions to the proposed quests. The engagement of players is not concentrated in a short period of time in which the game should react to specific emotions, i.e., student should not shoot monsters and increase their adrenaline like in a commercial game. As a consequence, we focused on the creation of a trusted relationship between students and their virtual tutors, the IPAs: “social intelligence in pedagogical agents may be important not just to gain user acceptance, but also to increase the effectiveness of the agent as a pedagogical intervention” has reported by Shen et al., in their study [95].

In order to understand students’ emotions, we preferred to start from the bottom, asking students to investigate their feeling in school context. We conducted a pilot study with an experimental group of students to collect a set of emotions and expected behaviour of the IPA. We believed that the strength of this approach is the collection of emotions’ coming directly from students/stakeholders, reducing the risk to model the IPA in a way that could be not in line with their expectations.

Figure 4.13. Emotions & Schooling pilot study

With an experimental group of students, we have worked on emotions, starting to analyse what they feel during real school scenarios, like a test session or an oral presentation. The research on emotion and schooling has been conducted with a high school class of 20 students aged 15 years. The research has been assisted by
two teachers: one from literature and one from computer science. We had seven meetings with the students, the project plan is reported in Table 4.2.

Table 4.2. Emotion and schooling project plan

<table>
<thead>
<tr>
<th>Single class schedule</th>
<th>110’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students</td>
<td>20</td>
</tr>
<tr>
<td>Number of lessons</td>
<td>7</td>
</tr>
<tr>
<td>School Grade</td>
<td>8th (15-16 years old)</td>
</tr>
<tr>
<td>School Subjects</td>
<td>Literature, Computer Science</td>
</tr>
<tr>
<td>Project timeslot</td>
<td>October-December 2014</td>
</tr>
</tbody>
</table>

Figure 4.14 reports the simple worksheet that students used during each lesson: they filled a specific column of the worksheet with the individual lesson outputs. At the end of the project, we used the information in the worksheets to collect data about emotions and VT wished behaviour.

In Appendix A, we elaborated a detailed summary of each lesson’s activities, reported in Table A.1.

Figure 4.15 reports some of the support materials that have been used for the research study. In particular we used the basic emotions schema, reported in 4.15(a) and taken from Lane & Nadel ([57]), to introduce students to emotions understanding. In order to analyse face expressions corresponding to certain emotions we used the famous TV series *Lie to me*, Figure 4.15(b). Finally, in Figure 4.15(c), we report a screen-shot of the Crazy Talk tool that we used to animate the VT.

Results of the pilot study are reported in Chapter 6, Section 6.2.
Once the pilot study finished, we collected all the results, populated the DB and start modelling the VT that embody the IPA. To prove the goodness of the IPA modelling, we built a stand-alone application in Android that makes use of the IPA as an Intelligent Virtual Tutor. The application and the settled experiments are fully described in Chapter 6, Section 6.2. We used the results to improve the IPA design and implementation.

In order to design the emotional behaviour of the IPA, we start working on players’ emotions detection in real time. We adopted two strategies.

The first strategy was to detect players’ emotions performing a Sentiment Analysis on player chat with the NPC, aka the IPA, because as explained by Calvo & Mac Kim in [12], Sentiment Analysis is able to catch players’ emotions in text. Depending on the captured emotion, the NPC reacts in different ways, following the results that came from the experimental study.
4.2 IPA

The second strategy that we put in place was to analyse the use of emoticons in chat messages among team members. Emoticons’ detection has not been used to model affect behaviour of IPA but to model the pedagogical intervention of the IPA, namely the unsolicited hints mechanism, as described in the following section.

4.2.3 Unsolicited Hints

An Intelligent Virtual Tutor (IVT) is able to answer the questions asked, in natural language, by the student/player. However, in our research model, the IVT becomes an IPA when it is able to act autonomously inside the game and provide unsolicited hints to the player. In fact, the IPA evaluates some parameters inside the game and decides when to provide help. The rational behind, is to keep the player involved in the game and raise his attention and motivation.

As explained by Parsons et al. in [80], an agent operating in a complex environment is inherently uncertain about that environment; it simply does not have enough information about the environment to know either the precise current state of its environments, nor how that environment will evolve. Thus, for every variable which captures some aspect of the current state of the environment, all the agent typically knows is that each possible value $x_{ij}$ of each $X_i$ has some probability $Pr(x_{ij})$ of being the current value of $X_i$. Writing $x$ for the set of all $x_{ij}$, we have:

$$Pr : x \in \mathbb{R} \mapsto [0, 1]$$

and

$$\sum_j Pr(x_{ij}) = 1$$

In other words, the probability $Pr(x_{ij})$ is a number between 0 and 1 and the sum of the probabilities of all the possible values of $X_i$ is 1. If $X_i$ is known to have value $x_{ij}$ then $Pr(x_{ij})=1$ and if it is known not to have value $x_{ij}$ then $Pr(x_{ij})=0$. Given two of these variables, $X_1$ and $X_2$, then the probabilities of the various values of $X_1$ and $X_2$ may be related to one another. If they are not related, a case we distinguish referring to $X_1$ and $X_2$ as being independent, then for any two values $x_{1i}$ and $x_{2j}$, we have:

$$Pr(x_{1i} \land x_{2j}) = Pr(x_{1i})Pr(x_{2j})$$

If the variables are not independent, then:

$$Pr(x_{1i} \land x_{2j}) = Pr(x_{1i}|x_{2j})Pr(x_{2j})$$

where $Pr(x_{1i}|x_{2j})$ is the probability of $X_1$ having value $x_{1i}$ given that $X_2$ is known to take value $x_{2j}$. Such conditional probabilities capture the relationship between $X_1$ and $X_2$, representing, for instance, the fact that $x_{1i}$ becomes much more likely when $x_{2j}$ is known to be true. If we take the set of these $X_i$ of which the agent is aware, the set $X$, then for each pair of variables in $X$ we can establish whether the pair are independent or not. We can then build up a graph in which each node corresponds to a variable in $X$ and an arc joins two nodes if the variables
represented by those nodes are not independent of each other. The resulting graph is known as a Bayesian network, and the graphical structure provides a convenient computational framework in which to calculate the probabilities of interest to the agent. In general, the agent will have some set of variables whose values it can observe, and once these observations have been taken, will want to calculate the probabilities of the various values of some other set of variables.

In order to realise the pedagogical intervention of the IPA we decided to monitor two distinct aspects of players behaviour in the game. The first aspect is strictly related to players’ advancement in the game, the second aspect is related to the emotions felt by players. In the latter case, we monitor the putative sentiment experienced by players in the game. Those choices result in 5 independent variables, which deal with game progression and emotions.

In the following section, we first analyse the observable variables and then the Bayesian network that we have designed.

To analyse the progress in games’ sessions, beside time that is a key factor, we decided to check two other important variables. The first variable is the overall score obtained by the player which reveals different actions undertaken in the game by the player himself. In fact, the score deals with answering quizzes, collecting objects and solving tasks. The other important value is how many questions player has already asked to the IPA, that we classified as required help.

Recapping, we monitor the following game’s variable:

<table>
<thead>
<tr>
<th>Progresses’ variables in game’s session</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time</strong>: elapsed time since the beginning of the game</td>
</tr>
<tr>
<td><strong>Score</strong>: overall score obtained by the player</td>
</tr>
<tr>
<td><strong>Required help</strong>: questions asked by player to IPA</td>
</tr>
</tbody>
</table>

We tried to infer players’ emotions by analysing the conversation among them and their IPA and analysing the chat among them and other team’s members. In fact, emotions can be expressed while conversing with IPA: we have noticed (cfr Chapter 6, Section 6.2), that players express their emotions with interlocutory phrases like ’Thanks!’ or ’Hell!’ . Performing a sentiment analysis on players’ sentences we can detect emotions in real time. In chat message among team’s members we do not perform sentiment analysis, because usually text is too short, but we monitor emoticons’ usage. Users of electronic methods of communication have developed visual clues that are associated with emotional states in an attempt to state the emotion that their text represents. These have become known as smileys or emoticons and are glyphs constructed using the characters available on a standard keyboard, representing a facial expression of emotion. When the authors of an electronic communication uses an emoticon, they are effectively marking up their own text with an emotional state. This marked-up text can be used to train a sentiment classifier if we assume that a smile indicates generally positive text and a frown indicates generally negative text. The use of emoticons, is one way of overcoming the domain, topic and time problems in sentiment analysis as suggested by Read
in [84]. That is why we decided to use both emoticons and sentiment analysis. To facilitate emoticons usage, we decided to use a limited numbers of emoticons: eight as the basic emotions.

<table>
<thead>
<tr>
<th>Emotions’ variables in game’s session</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emoticons</strong>: use of emoticons in chat messages among team members</td>
</tr>
<tr>
<td><strong>Sentiment</strong>: output of the sentiment analysis performed on player/IPA dialogue</td>
</tr>
</tbody>
</table>

In our model, each one of the *progress* variables has a state that can be: low, medium or high. To set properly the interval for each variable, we have taken the maximum value of each variables’ state and then divided it into three parts. For example, in the case of time, the value indicates if the player is at the beginning of the game session, in the middle or at the end. Some value can be inferred during assessment (like average time session), some others are fixed (like maximum score). Then, to each state we assigned a probability, as explained at the beginning of this section.

Similarly, for emoticons and emotions we have decided to divide each node into three states: negative, neutral or positive. Sentiment analysis performed on player/IPA chat automatically assigns sentiment tags to each sentence. These tags and their percentage are used to classify the expressed sentiment as neutral, negative or positive.

For emoticons, we imagined the following scenario. During a chat, team’s members use emoticons. We assigned a value to each emoticon, a number ranging from -4 to +4 as reported in Table 4.3.

<table>
<thead>
<tr>
<th>Table 4.3. Emoticons’ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Emoticons</strong></td>
</tr>
<tr>
<td>acceptance</td>
</tr>
<tr>
<td>anger</td>
</tr>
<tr>
<td>anticipation</td>
</tr>
<tr>
<td>disgust</td>
</tr>
<tr>
<td>fear</td>
</tr>
<tr>
<td>joy</td>
</tr>
<tr>
<td>sadness</td>
</tr>
<tr>
<td>surprise</td>
</tr>
</tbody>
</table>

In order to classify the player’s sentiment as neutral, negative or positive, we
summed all the used emoticons. We used the following simple classification to classify sentiments:

\[
\begin{align*}
\text{emot} \leq -1 & \rightarrow Etag = \text{negative} \\
-1 \leq \text{emot} \leq +1 & \rightarrow Etag = \text{neutral} \\
\text{emot} \geq +1 & \rightarrow Etag = \text{negative}
\end{align*}
\]

where \text{emot} is the sum of emoticons values used in the chat and \text{Etag} is the sentiment tag for emoticons.

The two sentiment variables, if both present are combined using the logic reported in Table 4.4.

### Table 4.4. Emotions (Emo) vs Emoticons (Emot)

<table>
<thead>
<tr>
<th>Emo</th>
<th>Neg</th>
<th>Pos</th>
<th>Neu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neg</td>
<td>Neg</td>
<td>Neu</td>
<td>Neg</td>
</tr>
<tr>
<td>Pos</td>
<td>Neu</td>
<td>Pos</td>
<td>Pos</td>
</tr>
<tr>
<td>Neu</td>
<td>Neg</td>
<td>Pos</td>
<td>Neu</td>
</tr>
</tbody>
</table>

To combine all the observable variables we used the Bayesian Network represented in Figure 4.17. The green nodes are fathers and visible in the network, the orange nodes are not visible but predictable.

![Bayesian Network](image)

**Figure 4.17.** Unsolicited hint Bayesian Network variables nodes

In the left side of the above figure, combinations of probability bring to a stuck state in which the player progresses are not good; in the right side of the figure, combinations of probability bring to a state in which the player is experiencing a negative emotional state. Combining probability of stuck and negative states lead to
the probability of providing an unsolicited hint to the player. In our algorithm, we consider the 5 variables independent and we check the value of the variables every $dt$, so that:

$$Pr(time \land score \land help) = Pr(time)Pr(score)Pr(help)$$

and

$$Pr(emo \land emot) = Pr(emo)Pr(emot)$$

During the game session, the system checks the 5 variables’ values and calculate the probability of the Help node. The Algorithm that run on the client side is reported in 2, while the algorithm that we run on the server side is reported in 3. The client algorithm calls the server algorithm UnsHint (described as Algorithm 3) and provide the runtime parameters values.

**Algorithm 2 Unsolicited Hint client side**

**Input:** time = t, score = s, help = h, teamChat = tc, IPAChat = ic

1: for $t = t + \Delta_t$ do
2: UnsHint(t,s,h,tc,ic) = (hi,hint)
3: if hi = true then
4: provide hint;
5: else
6: return
7: end if
8: end for

$\Delta_t$ is the time increment that we define in the game session.
Algorithm 3 Unsolicited Hint server side

Input: \( t,s,h,tm,ic; \)

\[
\text{for } i = [0,5]; i++ \text{ do}
\]
\[
2: \quad \text{for all } a = \text{inputVar do}
\]
\[
4: \quad \text{if } (a = tm \mid ic) \text{ then}
\]
\[
6: \quad \text{else}
\]
\[
8: \quad \text{end if}
\]
\[
10: \quad \text{end for}
\]
\[
Pr(\text{stuck}) = \sum_{1}^{3}Pr_i
\]
\[
12: \quad Pr(\text{negative}) = \sum_{4}^{5}Pr_i
\]
\[
Pr(\text{hint}) = PrCalc(\text{stuck})PrCalc(\text{negative})
\]
\[
14: \quad \text{if } random > Pr(\text{hint}) \text{ then}
\]
\[
16: \quad \text{else}
\]
\[
18: \quad \text{end if}
\]

where function Emoticons, parse the chat among teams’ members and evaluate emoticons usage and function Sentiment analyse the sentiment expressed by the player in chatting with his/her IPA. PrCalc calculates the probability of the node given the input value; function ProvideHint takes an hint from the NPC’s database.
Chapter 5

Realizing NewEarth

In this chapter, we explain the single components implementation. We explain the technologies that we have used and the reasoning behind each choice. The structure of the chapter retraces the structure of the previous one, the design phase, in which we introduced the presentation, logic and data layers. The last section of the chapter describes the IPA implementation.

5.1 VLE

The VLE has been implemented using different technologies, depending on the specific component. Two components, the PVB and the 3DVG, have been realised using Unity 3D\(^1\), the PA instead has been realised as a web application. The rationale behind is that, while the PA should be accessible anywhere any-time, PVB and 3DVG should be accessed in class and run respectively on a multimedia whiteboard (that usually runs on a Window system) and on tablets. The logic tier offers a set of web services developed using REST, while the database has been implemented using MySQL.

5.1.1 Presentation tier: Teachers’ workspace

The teachers’ workspace consist of two different applications which, as already explained in Chapter 4, have different objectives. The Professor App is very simple

---

\(^1\) Unity is a cross-platform game engine developed by Unity Technologies (https://unity3d.com/) and used to develop video games for PC, consoles, mobile devices and websites. Unity is notable for its ability to target games to multiple platforms. Within a project, developers have control over delivery to mobile devices, web browsers, desktops, and consoles.
and we didn’t implemented a *gamified* interface, while the Professor Virtual Board offers a *gamified* interface.

**Professor App**

The professor application is a simple web based app implemented using Ajax. Ajax is a set of web development techniques using many web technologies on the client-side to create asynchronous Web applications. With Ajax, web applications can send data to and retrieve from a server asynchronously (in the background) without interfering with the display and behaviour of the existing page. By decoupling the data interchange layer from the presentation layer, Ajax allows for web pages to change content dynamically without the need to reload the entire page. Ajax is not a technology, but a group of technologies. We used HTML and CSS in combination to mark up and style information. The DOM (Document Object Model) is accessed with JavaScript to dynamically display - and allow the user to interact with - the information presented. Finally we used JavaScript and the XMLHttpRequest for exchanging data asynchronously between the browser and the server to avoid full page reloads. The main features of the web app have been already presented in Section 4.1.1. Once registered/login to the web app, teachers can add or remove class. Figure 5.1 shows two screen-shots of PA’s view.

![a) PA login view](image)

![b) PA class selection view](image)

**Figure 5.1.** PA login and class

The most important operation teachers can do is to start/stop sessions. When a session is started a unique PIN is generated, see Figure 5.2(b).
The PA is continuously updated in order to show live score of an active game session. Usually professors follow the game from the PVB, however they can also see results from the PA, as represented in Figure 5.3. These data are persistent and can be visualised at any moment: professors can check score, that means advancement in the game, whenever they want. They can analyse game progression and results to plan intervention for future sessions or simply to assess players’ performance.

**Professor Virtual Board**

PVB has been implemented using Unity 3D mainly for two reasons. The first one is that the PVB should be appealing and offers an interface similar to the 3DVG one, which has been developed in Unity as well. The graphic aspects are important in virtual SGs and using Unity was easier to make the PVB and 3DVG graphically...
compliant. Secondly, due mainly to time and resource constraints, we decide to use the same technology for two components that should be strictly related and communicate all along games’ sessions.

As explained in Section 4.1.1, the teacher launches the PVB interface and starts a game session using the PIN generated by the PA. So, the first step is to pass the PIN through the interface and use it as a login identifier, as shown in Figure 5.4. Thus, professors do not use their credentials to start the game session but just the PIN.

Figure 5.4. PVB login interface

The PVB board is divided into several windows, visible at the same time, see Figure 5.5. In the upper part of the board there are three windows:

- First window on the left: list of teams
- Central window: List of players of the selected team
- Right window: Knowledge Forum of the selected team

In the bottom part of the board there is a 2D representation map of the 3DVG environment. There is a spaceship and the three available planets: Maya, Melissa and Violet. Real positions of players in the game are visualised in the 2D map. In this way, it is possible to localise players in the spaceship.

Once a professor is logged into PVB, if it is the first time he/she should add new teams. The players will appear just when they register and hook to a specific team. This mechanism implies a coordination between teacher and students. In fact, once the class has been divided into teams and the teacher has created the virtual teams from the PVB, the students register using the 3DVG, hook to the chosen team, select a role and become player. At this stage, they appear on the PVB.
From the PVB, professors can (See Figure 5.6):

- See single player info (score, role, assigned task)
- See players’ position on the 2D map
- Validate or not players’ notes

Beside the individual score, professors can see also the team score. The list of teams appears following a decreasing score order.

The 3DVG updates continuously the database, calling the web services offered by the logic tier, writing info about players.
As mentioned before, players can share knowledge notes about the game content and their discoveries in the game. The Knowledge Forum is visible in the right top side of the PVB. When notes are sent by players from their 3DVG, they appear in the PVB. Then professor has two possibilities, validate or not the note. If the professor evaluates that the note make sense for the game and it is useful for the community he validates it. In this case, the score increases and it is automatically updated in the PVB; the note is added to the DB and all team members can see the note in their 3DVG. If professor decides to reject the note (Figure 5.7), then the note is coloured in red and it will not appear in the players’ 3DVG notes list.

Unity, beside the possibility to model 2D or 3D environment, offers the possibility to add scripting. Scripting is an essential ingredient in all games. Even the simplest game needs scripts, to respond to input from the player and arrange for events in the game-play to happen when they should. In our implementation we have used scripting in many ways for different purposes. We used scripts to create graphical effects and to control the physical behaviour of objects. But mainly we used scripting to communicate with web services offered by the logic-tier. In fact, the majority of the actions in the PVB interface, like login, add teams, manage notes, etc. are done executing scripts that calls our corresponding web services.
Following the schema represented in Figure 5.8 we were able to decouple the game logic from the game interface. We used C# to implement the scripts where the web services are called. The same considerations are true for the 3DGV described in the following Section.

5.1.2 Presentation tier: Students’ workspace

The students’ workspace is formed, as explained in paragraph 4.1.2, by Opedia® and the 3DGV.

The 3DGV component has been implemented using Unity 3D which, as already explained in the previous paragraph, is a free game engine particularly valuable for designing 3D virtual games. In the implementation phase, the main challenges have been to:

- Design and build from scratch a 3D virtual environment that was in line with the plot
- Model the three exo-planets simulation rooms
- Optimize the game for mobile

Even if the game is a RPG we have decided to not implement it as a multi players game. That means that we keep roles but player cannot physically meet in the 3D environment. They can see each others’ positions on 2D map in the PVB and they can chat. That decision was taken to keep the implementation simple and especially to keep the network load lower. In fact, usually network’s connection in schools does not support a big workload. For the same reason, we tried to reduce the communication between the 3DVG and the logic tier at minimum and keep the game self-consistent as much as possible. In particular, the 3DVG makes use of:

- login/register services when the player authenticates
- players’ services for communicating updates like score, position, etc.
- knowledge services when players want to share a note
- IPA’s services when players want to ask a question to the IPA
- AI algorithm to handle unsolicited hints

All these services lie in the logic tier. We implemented several C# scripts that make the external calls and handle game’s variables.

Concerning the 3DVG usage players have to register the first time they want to access the game. They have to provide some personal data, choose a team and a role. Then, in the following accesses they can simply login; Figure 5.8 shows the login interface.
Once the player has logged in, he can start playing. Players can explore the 3D virtual environment (Figure 5.10) that is formed by a spaceship and three exo-planets simulation rooms.

Figure 5.10(d) shows a stone, that is one of the six objects that the player has to collect. Objects are needed to solve the assigned tasks and change according to the role and to the experiment. For example, the physicist has to collect a stone and a goniometer to solve the experiment about the inclined plane. An interactive screen and a gate are also visible in the picture. Players should reply to quizzes (via interactive screen) in order to access the simulation room (via a tele-transport gate).

All the functions that players can use are also visible in the right top part of the figure, from left there are:

- Button to insert task solutions
- Button to open Opedia®
- Button to dialogue with IPA
- Button to add a note

The inventory of the collected objects is also visible in the left top part of the figure.
5.1 VLE

(a) Central deck of the spaceship  
(b) Corridor that bridges different spaceship rooms  
(c) Trophy room  
(d) Planet entrance  

Figure 5.10. View of the 3D virtual environment

As already explained, the interactive screen gives access to quiz panels (Figure 5.11(a)) which serves to gain access to the planet’s simulation room.

(a) Example of quiz to be solved to gain access to the planet 3D environment  
(b) Player answered correctly to a quiz  

Figure 5.11. Example of quiz panels

Figures 5.12 and 5.13 show respectively the panels that we have implemented to dialogue with the IPA and to add a note.
As better explained in Section 5.2, the IPA has been implemented as a service. In this way, it has been easily integrated into the 3DVG as all the other web services, following the logic in Figure 5.8.

Another function that the players can use during a game session is the chat among team members. Chat has been realised using an off the shelf chat system (developed as a service), that offers a free of charge API that has been easily integrated into Unity. In the chat panel (Figure 5.14), we have explicitly represented the 8 basic emotions to encourage players to use them, fostering the idea of emoticons for detecting emotions in game.

\footnote{https://www.photonengine.com/en/Chat}
A very important aspect of the 3DVG is the content and the players’ assigned tasks. As already mentioned, each player (role) should solve three different tasks, each focused at studying a specific characteristic of the proposed exo-planets. Problems have been proposed by field experts and cover different STEM subjects taught in schools. The exo-planets have been modelled by field experts and their data are perfectly consistent with Physics models. As an example we report a screen-shot of the 3D exo-planet Maya, in Figure 5.15.

The assigned tasks are proposed as experiments or simulations in which players can use interactive panels. Players should solve problems and then provide their solutions. Some of the interactive panels are reported in Figure 5.16.
The individual game experience through the 3DVB ends when the player has solved all the assigned tasks and has provided the final solutions. The SG ends when teams have produced a detailed report on their conclusion about which is the best exo-planet to inhabit.

5.1.3 Logic Tier

The logic tier hosts, in our implementation, all the services that rule users’ actions and data. Users, as already explained, can be teachers, students, and players. Students become players when they register for the first time. All the services have been implemented as web services using REST style, as fully explained in the following section.

Web Services

Representational state transfer (REST) or RESTful Web services are one way of providing interoperability between computer systems on the Internet. Representational State Transfer (REST) is an architectural style that specifies constraints, such as the uniform interface, that if applied to a web service induce desirable properties, such as performance, scalability, and modifiability, that enable services to work best on the Web. In the REST architectural style, data and functionality are considered resources and are accessed using Uniform Resource Identifiers (URIs), typically links on the Web. The resources are acted upon by using a set of simple, well-defined operations. The REST architectural style constrains an architecture to a client/server architecture and is designed to use a stateless communication protocol, typically HTTP. Usually, clients and servers exchange representations of resources by using a standardised interface and protocol.

REST-compliant Web services allow requesting systems to access and manipulate textual representations of Web resources using a uniform and predefined set of stateless operations. In our implementation, we have implemented several web services, divided in 8 macro domains, which are:

- Authentication
- Class
We used the Swagger Editor\(^3\), which is an open source editor to design, define and document RESTful APIs in the Swagger Specification. The Swagger specification is a definition format to describe RESTful APIs. It creates a RESTful interface for developing and consuming an API by effectively mapping all the resources and operations associated with it (see also Figure 5.17).

![Swagger Editor Screenshot](http://swagger.io/)

**Figure 5.17.** Screen-shot of web services swagger documentation

All the operations bound to the activities related to security issues are under the umbrella of *authentication*. We have implemented a mechanism of security token. Security tokens are used to prove one’s identity electronically. Generally, the token is used in addition to or in place of a password to prove that the *customer* is who he/she claims to be. The token acts like an electronic key to access something. In our case, a token is used by applications (namely PA, PVB, 3DVG) that want to call the available web services. A unique token is generated and it has a fixed duration (like 15 minutes). Each time a web service is called, if the token is expired, it has to be newly generated.

**Authentication web services**

- **Type**: post  |  **Name**: `authentication`
- **Type**: post  |  **Name**: `authentication/pin`

\(^3\)http://swagger.io/
The class web services are related to all the operations that treat the class as a whole: the class is a group of students which attends a specific course and taught by a group of teachers.

### Class web services

1. **Type:** post | **Name:** class
2. **Type:** delete | **Name:** class/{id}
3. **Type:** post | **Name:** class/{id}/startSession
4. **Type:** post | **Name:** class/{id}/stopSession
5. **Type:** get | **Name:** class/{id}/teams

The concept of class is also related to the concept of teams. In fact, each class is divided in teams, which are sub groups of students that play together.

The web services under the umbrella of professor are services used to authenticate the professor, to remove him/her from the users’ list or to associate him/her with a given class.

### Professor web services

1. **Type:** post | **Name:** professor
2. **Type:** delete | **Name:** professor/{id}
3. **Type:** get | **Name:** professor/{id}
4. **Type:** put | **Name:** professor/{id}
5. **Type:** delete | **Name:** professor/{id}/teaching/class/{class_id}
6. **Type:** post | **Name:** professor/{id}/teaching/class/{class_id}

The web services under the umbrella of note are the ones related to the possibility for players to add notes to the Knowledge Forum. Then, teachers can approve or not the note sent by the player.

### Note web services

1. **Type:** delete | **Name:** note
2. **Type:** get | **Name:** note/{id}
3. **Type:** put | **Name:** note/{id}/assessment

The services in the school area are useful to recover information about available schools and classes for each school. To have the list of schools we have used the unique key provided by the government to identify Italian schools.
School web services

1. **Type**: get | **Name**: school/list_All
2. **Type**: get | **Name**: school/{processing_number}/classes

The web services under the umbrella of student deal mainly with recovering information about players during games’ sessions. In particular, we should know players’ position, score, notes, quiz, task, hint, etc.

Student web services

1. **Type**: post | **Name**: student
2. **Type**: delete | **Name**: student/{id}
3. **Type**: get | **Name**: student/{id}
4. **Type**: post | **Name**: student/{id}/gameProgress
5. **Type**: put | **Name**: student/{id}/gameProgress
6. **Type**: delete | **Name**: student/{id}/membership/team/{team_id}
7. **Type**: post | **Name**: student/{id}/membership/team/{team_id}
8. **Type**: post | **Name**: student/{id}/message
9. **Type**: get | **Name**: student/{id}/messages
10. **Type**: post | **Name**: student/{id}/note
11. **Type**: post | **Name**: student/{id}/position
12. **Type**: put | **Name**: student/{id}/position
13. **Type**: post | **Name**: student/{id}/question
14. **Type**: post | **Name**: student/{id}/quiz
15. **Type**: post | **Name**: student/{id}/score
16. **Type**: put | **Name**: student/{id}/score
17. **Type**: post | **Name**: student/{id}/unsolicitedHint

Finally, we have the team web services that are used by the system to recover information organised by teams’ view.
We now provide an explanatory example of how the web services are called in our system. In Figure 5.18, we report a state diagram (5.18(a)) in which a student adds a note. Figure 5.18(b) shows which web services are called when player sends the note.

![State diagram](image)

**Figure 5.18.** Player adds a note from 3DVG interface

Figure 5.19 shows what happens when the note appears in the PVB and the teacher validates it or not. If the assessment is positive, the note is added to the DB and the student’s score is updated. If the assessment is negative, the note is deleted. We have chosen this example because we want to show how all the interfaces are related and how they work using the web services in the logic tier.
5.1 VLE

(a) State diagram

(b) Web services diagram

Figure 5.19. Teacher validates/does not validate a note from PVB

5.1.4 Data Tier

We have implemented the database starting from the E-R diagram reported in Section 4.1.4. Hereafter we report the logical translation of the diagram. Each entry corresponds to a table and each attributes (in brackets) equals a column. Underlined attributes are the primary key link for the different tables; the other possible links are reported under each entry.

Logic and data tier are strictly connected since many of the web services in the logic layer, can access the database both for reading and writing data. The presentation tier is not directly connected to the DB and every data information should pass through the logic tier, as depicted in the overall architecture, see Figure 4.2.
As already mentioned, one of the novelties of the IPA has been its implementation according to the as-a-service paradigm. All the Natural Language Processing (NLP) related computations are performed on the server side, while the client side represents a consumer application that interfaces with our system, in the specific case the
3DVG. In the following section, we provide a detailed analysis of the modules that compose the IPA. We can distinguish different levels, defined below.

**NLP Logic.** The NLP Logic module deals with the elaboration of the input user’s sentence and the subsequent analysis of the data available in the NPC KDB in order to identify the best answer. Moreover, it tracks the evolution of the currently active conversations. The computation of the sentences happens on two different logical levels, which correspond to two sub modules, one high and strategic level and one low and tactical level.

**Strategic Level.** The Strategic Level module deals with the high level sentence computation and the management of the conversations. For each application, that requested the system’s services it tracks how much NPC knowledge the users discovered by asking their questions, thus calculating a score for each active conversation. The score assignment mechanism has critical importance during gameplay since it provides feedback to the user about how the conversation is evolving and about the effectiveness of his exploration of the NPC’s knowledge space, which is unknown when the game starts.

**Tactical Level.** The Tactical Level module performs a low level sentence processing, not taking care of high level details, like management of the sentences or the conversation’s monitoring activity, for which the Strategic Level is in charge. It takes a sentence as input and it computes it by analyzing each word, isolating from the database data belonging to that NPC’s KDB and trying to identify, among the available ones, an answer that best matches the user’s question by using the text classification algorithms implemented by the system. This module represents the core of the NLP computation.

To implement the NLP computation we used Stanford CoreNLP\(^4\). Stanford CoreNLP provides a set of natural language analysis tools. It can give the base forms of words, their parts of speech, whether they are names of companies, people, etc., normalise dates, times, and numeric quantities, mark up the structure of sentences in terms of phrases and word dependencies, indicate which noun phrases refer to the same entities, indicate sentiment, extract particular or open-class relations between entity mentions, get quotes people said, etc.

**Database.** The database stores all the knowledge defined by the authors. It stores a set of registered applications and a set of all the authors, as well as a set of all the registered users. Of critical importance is the storage of all the statistical parameters relative to the preprocessing of the KDBs’ sentences. Furthermore, in the database, the system keeps track of the history of all the conversations that took place, which can be helpful for further analysis of the users’ data in order to improve the algorithms and also the set of answers.

We have already seen how templates are stored in the NPC Database in Section 4.1.4, Figure 4.12.

To model the avatars that embody the IPA we used the CrazyTalk tool. This commercial tool has also a Unity plugin which allows to import avatars animations in Unity scene. Figure\(^2\) 5.21 presents the full implementation process from avatar creation, voice lip-sync generation and talking script editing in CrazyTalk to Unity

\(^{4}\text{http://stanfordnlp.github.io/CoreNLP/}\)

\(^{2}\text{Ri-edited from https://www.reallusion.com/crazytalk7/unity-features.aspx}\)
export and integration.

Using the CrazyTalk plugin we were able to have animated avatars that embody the IPA in game scene. In particular we use it to introduce the assigned task to the player, like in Figure 5.22.

To perform sentiment analysis on players sentences during conversation with IPA, we used IBM Bluemix Watson\(^5\). We used the API called Tone Analyzer, which uses a linguistic analysis to detect three types of tones in written text: emotions, social tendencies, and writing style. We uses the Tone Analyzer service to understand emotional context of conversations between IPA and players.

From client side, we implemented a function that was executed every 3 minutes (Algorithm 2). It collects the needed variables’ values, uploads them to the server and checks the latter’s response. A boolean value (coming from the algorithm that we implemented and described in Algorithm 3) indicates if the player needs an unsolicited hint; if so, a field containing the hint (taken randomly from the NPC

---

Feedbacks are collected to evaluate players opinion on hints mechanism, as explained in Chapter 6, Section 6.5.
Chapter 6

Game evaluation & learner assessment

In this chapter we describe the three test runs that were conducted in a formal learning environment. The first test run involved roughly 50 students and aimed at testing the IPA in a stand alone Android application. The second and third test runs aimed at testing NewEarth. The outcome of the second run, in which 42 students were involved, has been used to improve NewEarth. The last run (47 students) gave us the opportunity to test also the so called 21st Century skills.

6.1 Introduction

In this chapter we pose the attention on two important aspects of New Earth evaluation. In one hand we measured the game effectiveness and its goodness in achieving the prefixed goals: namely to help students to learn STEM in a playful way and acquire the 21st century skills. We evaluated technical choices and how they sustained or not the pedagogical goals. On the other hand, we had to effectively measure if learners learned while playing and thus we performed students/player assessment.

Assessment describes the process of using data to demonstrate that stated learning goals and objectives are actually being met. Michael & Chen [70] state 'Serious games, like every other tool of education, must be able to show that the necessary learning has occurred.' SGs, to be considered a viable educational tool, must provide some means of testing and progress tracking; testing must be recognizable within the context of the education or training they are attempting to impart. However, learning is a complex construct, difficult to measure, and determining whether a simulation or SG is effective at achieving the intended learning goals is a complex,
time consuming, expensive, and difficult process (see Hays in [44] and Enfield et al. in [32]).

Generally speaking, an assessment can be described as either (i) *summative* whereby it is conducted at the end of a learning process and tests the overall achievements, or (ii) *formative* whereby it is implemented and presented throughout the entire learning process and continuously monitors progress and failures, (Boston in [9]).

Considering the specific SG domain, Michael and Chen describe three primary types of assessment: (i) *completion assessment*, (ii) *in-process assessment*, and (iii) *teacher assessment*. The first two correspond to summative and formative assessments, respectively. Completion assessment is concerned with whether the player successfully completes the game. In a traditional teaching environment, this is equivalent to asking, "Did the student get the right answer?". Teacher assessment focuses on the instructor’s observations and judgements of the students while they are playing the game and typically aims at evaluating those factors that the functionalities/logic of the game are not able to capture.

The most common method of post assessment currently consists in testing player’s knowledge about what they learned by way of a survey/test/questionnaire or teacher evaluation. This method is frequently employed because it is the simplest to implement, but it relies on the opinions of the players and does not depend on all the information that can be collected regarding what happened within the game.

In order to test NE we have adopted different strategies. The first assessment, namely run no.1, has been conducted to test the effectiveness of the Virtual Tutor and of the Dialogue Management System developed by us. We developed a stand alone application, an Android app available on Google Play store, and test it in context of formal education. The app design and implementation are reported in Appendix A, while the assessment is reported in Section 6.2. The first test run gave us a preliminary validation of some functionalities of the IPA from both an engineering and a pedagogical perspective.

The second assessment has been done on the first version of NE. This version of the SG was almost complete, except of few implementations still missing. We evaluated both summative and formative learning, playing the game for four lessons in a formal learning environment. The second test run is described in Section 6.3.

The third assessment has been done on the final version of NE. This version of the SG was complete, including lots of improvements. We tested this version with 2 classes, playing the game for five lessons, in a formal learning environment. The full description of this third test run is reported in Section 6.4.

### 6.2 Ecological test run no. 1

#### 6.2.1 EIVT and Android app

Following an Iterative Model approach to the system design and implementation, we decided to first implement the Emotional Intelligent Virtual Tutor (EIVT), as a precursor of the IPA, and test it in a real school scenario. Moreover, as already
explained in 4.2.2, we conducted a pilot study with students in order to investigate emotions and schooling, and to model the EIVT consequently. The research was conducted following the steps listed below:

- Conduct a pilot study with high school students to understand students’ emotions in school environment
- Gather wished behaviour of the Virtual Tutor
- Translate wished behaviour in an Emotionally Intelligent VT
- Implement the EIVT as a service
- Implement a Dialogue Management System, taking into account both learning and emotional aspects
- Implement animated avatars that embody the EIVT
- Implement an app to exploit the EIVT
- Test the app in schools to verify the goodness of the EIVT

The design and development of the EIVT have been done following the research model reported in Figure 6.1.

![Figure 6.1. EIVT research model schema](image-url)
6.2 Ecological test run no. 1

We worked with three different groups of students coming from three different schools, that guaranteed a more heterogeneous sample of students. The first class (Group1 in the model) has been selected to work with us from the beginning of the research. In fact, we worked with this group to model the emotional behaviour of the tutor conducting a three months pilot study (as described in Section 4.2.2), together with Literature and Computer Science teachers.

Once the data from the pilot study was collected and analysed, we built the EIVT implementing the Artificial Intelligence algorithm described in Section 4.2.1. We performed a first set of tests (after roughly one-year development) with Group1 and Group2 mainly aimed at getting feedback about liking and usability through the use of questionnaires. The results of the first two test runs, beside the technical validation of the AI algorithm performances, were used to optimize the application and to inquire how students perceive the emotional behaviour of the VT. The third and last test run has been conducted in a significantly different way to mainly assess summative outcomes. We divided the class into an experimental and control group to verify the validity of the EIVT for learning purposes. Subjects of study have been provided by teachers as possible learning insights, in fact teachers were able to upload their content using an authoring tool.

A detailed description of the Android app (see Figure 6.2) can be found in Appendix A.

6.2.2 Ecological Test

Tests have been conducted in a real context of formal education in order to assess usability and usefulness of the EIVT, plus the potential benefit of an emotionally intelligent EIVT. Some simple persistence learning outcomes have been assessed comparing control and experimental group after one month from the first test run. Tests were also used to verify the goodness of the AI algorithm and threshold adjustments.

The Android application came with two different learning paths: one in STEM (the terrestrial atmosphere) and one in humanities (the New Deal). These two learning paths have been chosen and written by teachers and have been identified as possible insights with reference to the traditional study plan of students aged 16-18.
6.2 Ecological test run no. 1

years. Before proposing the identified insights, we investigated with teachers of the three schools that students were not acquainted with these topics. They just had a simple conversation with the students about those topics and, because students were really ignorant about the content, we decided to not perform preliminary tests. In the first two runs, each student was free to choose the preferred topic. Students were also able to choose one of the two available animated avatar: a scientist, designed like a cartoon (shown in Figure 6.3(a)) or Anne, a young woman (shown in Figure 6.3(b)).

![Scientist](a) Scientist ![Anne](b) Anne

Figure 6.3. CrazyTalk animate avatar used to embody EIVT

Students were able to play with the application for approximately 40 minutes: we asked them to play with the application asking questions to the EIVT concerning the subject of study and trying to figure out the correct answer to the initial quest. Once students had finished asking questions, they compiled a questionnaire structured in three main sections: appreciation, usability and content appropriateness. The first run of test validation occurred on the 19 of November 2015 (Group1: 11 students, all males, aged 17-18). It is extremely sad noting that after one year, Group1 diminishes from 20 students to 13 (just 11 present the testing day) due to scholastic dispersion. The second run occurred on the 20 of January 2016 (16 students, two females, aged 17-18). The third run occurred on the 8 of February 2016 (13 students, 8 males and 5 females, aged 17-18). A summary is reported in which students were able to answer the questionnaire choosing among 5 ordered options from *Yes a lot* to *Not at all*. We assigned a score of 5 to *Yes a lot* and a score of 1 to *Not at all*. 
Table 6.1. Ecological test sessions in formal learning environment

<table>
<thead>
<tr>
<th>Date</th>
<th>Number of students</th>
<th>Age</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>19/11/2015</td>
<td>11 (all ♂)</td>
<td>17-18</td>
<td>Play 40 minutes, questionnaire</td>
</tr>
<tr>
<td>20/01/2016</td>
<td>16 (14 ♂, 2 ♀)</td>
<td>17-18</td>
<td>Play 40 minutes, questionnaire</td>
</tr>
<tr>
<td>08/02/2016</td>
<td>13 (8 ♂ and 5 ♀)</td>
<td>17-18</td>
<td>Play 40 minutes, read texts, control and experimental group, written test, questionnaire</td>
</tr>
<tr>
<td>09/03/2016</td>
<td>13 (8 ♂ and 5 ♀)</td>
<td>17-18</td>
<td>written test</td>
</tr>
</tbody>
</table>

The appreciation section of the questionnaire aimed at understanding if students liked the idea of a personalised learning path and the possibility to ask questions in natural language to the EIVT. Table 6.2 reports the questions asked and the results of the three runs.

Table 6.2. Appreciation’s section, question and results

<table>
<thead>
<tr>
<th>Questions</th>
<th>Average and SD 1-2</th>
<th>3 run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you think free questions are helpful?</td>
<td>(3.9 ± 0.6)</td>
<td>(4.1 ± 0.4)</td>
</tr>
<tr>
<td>Is it desirable to have a personalised learning path?</td>
<td>(4.0 ± 0.6)</td>
<td>(4.1 ± 0.5)</td>
</tr>
<tr>
<td>Can EIVT help while studying?</td>
<td>(3.8 ± 0.9)</td>
<td>(4.2 ± 0.4)</td>
</tr>
<tr>
<td>Is it important to have an empathic guide?</td>
<td>(3.4 ± 1.0)</td>
<td>(4.1 ± 0.7)</td>
</tr>
<tr>
<td>Would you like to chat about personal matters?</td>
<td>(3.4 ± 1.0)</td>
<td>(3.6 ± 1.0)</td>
</tr>
<tr>
<td>Did you perceive the application as a game?</td>
<td>(2.8 ± 1.1)</td>
<td>(2.2 ± 1.1)</td>
</tr>
<tr>
<td>Did you enjoy playing?</td>
<td>(3.4 ± 1.1)</td>
<td>(3.0 ± 1.1)</td>
</tr>
<tr>
<td>Would you recommend the app to your classmates?</td>
<td>(3.7 ± 0.9)</td>
<td>(3.6 ± 0.7)</td>
</tr>
</tbody>
</table>

The usability section of the questionnaire aimed at understanding if the application was clear enough and easy to use. In addition, we asked if the proposed avatars were endearing or not. Questions and results are reported in Table 6.3.
6.2 Ecological test run no. 1

Table 6.3. Usability section, questions and results

<table>
<thead>
<tr>
<th>Questions</th>
<th>Average 1-2 runs</th>
<th>3 run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the app interface clear?</td>
<td>3.5 ± 1.0</td>
<td>3.6 ± 0.8</td>
</tr>
<tr>
<td>Is the app usage clear?</td>
<td>3.7 ± 1.0</td>
<td>3.8 ± 0.7</td>
</tr>
<tr>
<td>Is the animated avatar endearing?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anne</td>
<td>3.5 ± 0.7</td>
<td>1.5 ± 0.6</td>
</tr>
<tr>
<td>Scientist</td>
<td>3.7 ± 0.6</td>
<td>2.4 ± 1.0</td>
</tr>
<tr>
<td>Are avatar verbal responses clear enough?</td>
<td>3.6 ± 0.8</td>
<td>3.1 ± 0.7</td>
</tr>
<tr>
<td>Did you consider the response of avatar to be affective?</td>
<td>3.2 ± 1.1</td>
<td>2.8 ± 0.8</td>
</tr>
</tbody>
</table>

The content appropriateness section aimed at verify if the provided EIVT answers were clear and if students had the impression that the answers replied by the EIVT were consistent with the questions asked.

Table 6.4. Content’s section, questions and results

<table>
<thead>
<tr>
<th>Questions</th>
<th>Average 1-2 runs</th>
<th>3 run</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are explanations clear enough?</td>
<td>3.8 ± 0.8</td>
<td>3.5 ± 1.2</td>
</tr>
<tr>
<td>Are answers consistent with the questions asked?</td>
<td>3.9 ± 0.8</td>
<td>3.6 ± 1.2</td>
</tr>
<tr>
<td>Are tips useful?</td>
<td>3.2 ± 0.9</td>
<td>3.5 ± 1.1</td>
</tr>
</tbody>
</table>

Table 6.5. Questions asked

<table>
<thead>
<tr>
<th>Questions</th>
<th>Answers</th>
<th>Number of students 1-2 runs</th>
<th>Number of students 3 run</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many questions did you ask before quitting the game?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Almost 5</td>
<td>12</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Almost 10</td>
<td>9</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Almost 20</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>More than 10</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

In the questionnaire, as reported in Table 6.5, we asked students to indicate how many questions they asked to the VT before quitting the game. We also asked
students if they thought that the app was more effective for STEM or humanistic subjects.

As already mentioned, the third run has been done after the optimization of the algorithm and after adding a new feature to the application: the bar showing the percentage of explored knowledge. Beside the questionnaire the class has been asked to undertake a real school test about the subject of study. Because the third run aimed at verifying if the game can produce different results from the standard lesson, the class was divided into two groups: experimental and control. For the first 40 minutes, the control group studied the STEM topic in a traditional way by reading a document written by the domain expert. At the same time, the experimental group played with the app selecting the STEM topic as well. Then, the two groups swapped. The first group played with the app selecting the Humanities topic while the second group studied the Humanities topic by reading a document prepared by the domain expert. At the end of the entire session, we prepared a test session to verify the learning outcomes. The test session included 10 close questions (multiple choices) about the subjects of study. After 30 days we asked the same questions to the students to verify if the use of the EIVT improved persistent learning outcomes.

Table 6.6 reports what we have found in the third run, expressing the percentage of correct answers for the four groups.

<table>
<thead>
<tr>
<th>Topic</th>
<th>Experimental (%)</th>
<th>Control (%)</th>
<th>Exp. + 30 days (%)</th>
<th>Cont. + 30 days (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEM</td>
<td>77,1</td>
<td>37,1</td>
<td>68,6</td>
<td>25,7</td>
</tr>
<tr>
<td>Humanities</td>
<td>60</td>
<td>42,8</td>
<td>54,3</td>
<td>31,4</td>
</tr>
</tbody>
</table>

6.2.3 Discussion of Results

Data analysis, coming from the results of the appreciation section reported in Table 6.2, shows that students like the idea of a personalised learning path and the possibility to ask free questions guided by a Virtual Tutor: in fact 75% of the students liked this possibility. The aspect of the empathetic guide was felt as well: 68% of the students liked the idea to have an empathic guide and would like to chat also about personal matters. 68% of students enjoyed the app and would recommend it to others, even if just 50% of them perceived it as a game. The analysis of the results of the usability section, reported in Table 6.3, shows that 68% of students found the interface clear. Most of the interviewed disliked the scientist avatar (70%) while liked the young girl (72%) in the first and second run. In the third run the percentage of students liking the two avatars decrease even if the avatar were exactly the same. Because of that we decided to sketch completely new avatars for New Earth. Results in Table 6.4 revealed that 78% of the students thought avatar explanations were clear and declared that provided answers were in line with the questions. 64% declared that tips were useful. We also asked how many questions they asked the avatar before quitting the game (Table 6.5): 46% gave up after few questions, 34% ended the game after roughly 10 questions, 15%
around 20 and just 1 student declared to have asked more than 20 questions. Those results were obtained in the first two runs. This is probably due to the fact that very often, pertinent questions received no answer due to the very high threshold settings of the AI algorithm. However, after threshold calibration, in the third test we revealed that 38% ended the game after roughly 20 questions, while 30% gave up after few questions. So we had an increment of roughly 20% in content exploration. The average number of asked questions per student increased from the second to the third run. This result is probably due to the higher number of correct replies by the EIVT that encouraged students to keep asking. As a general comment, we have seen that standard deviation is quite high and that it is probably due to two major facts: the student sample was not very high and there is a big fluctuation in student perception of game and virtual interaction. This hypothesis is confirmed by the fact that just 50% of the students perceived the application as a game, even if the majority (64%) enjoyed the application and will recommend it to their classmates (68%). The knowledge exploration’s bar, added in the third run was not enough, students in fact complained about the lack of a score system. However, many students declared that seeing the percentage of explored knowledge was an incentive to ask more questions.

The results coming from the third run, even if conducted on a small sample, give us some clear indications about research objectives. As we can infer from Table 6.6, the percentage of correct answers was higher using our application respect to a more traditional approach (reading a text). In fact, roughly 30% more students (average between STEM and humanistic) replied correctly. It seems also that the app usage is more effective for STEM subject. Another interesting result is the one related to the persistence of the learning outcome. In fact, the question of retention and how SG supports or undermines long-term learning is still an open issue. From our test session there was a 40% more correct answers (in STEM) after 30 days in the experimental than in the control group. Even in history the percentage is encouraging assessing 23% more of correct answers.

The results go in the direction that when one-on-one natural language tutoring is compared to a less interactive control condition that covers the same content, then the tutees will learn more than the non-tutees. The results of the EIVT assessment were encouraging and we integrated the EIVT into the SG. The EIVT - in the game - became an IPA capable of pedagogical intervention via unsolicited hints.

In the next two sections we describe the two ecological tests that we conducted with High School students to test NE.

### 6.3 Ecological test run no. 2: First version of NE

In this section, we describe the ecological tests aimed at assessing the first version of NewEarth.

#### 6.3.1 Ecological environment description

In this test run, we worked with two classes both formed by 21 students from an Italian High School, as reported in Table 6.7. The two classes were in the same school but they differ in the specialization. While in the first two years of study...
they have similar subjects, usually the consumers are distributed differently in the
two specialization. Students that plan to continue studying after high school choose
Scientific Institute while students that plan to end studying after high school choose
the Industrial & Technical Institute.

Table 6.7. Test no. 2: students sample

<table>
<thead>
<tr>
<th>School Name</th>
<th>Classes</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITI Pascal</td>
<td>1F (Scientific Institute)</td>
<td>21 (9♀, 12♂)</td>
</tr>
<tr>
<td>ITI Pascal</td>
<td>1N (Industrial and Technical Institute)</td>
<td>21 (6♀, 15♂)</td>
</tr>
</tbody>
</table>

We had 4 lessons between November 2016 and January 2017. In Table 6.8 we report lessons schedule and plan. The 4th lesson was planned for 20 December 2017. Unfortunately that particular day students were not at school and we postponed the lesson after Christmas vacation. That resulted in a delay between the third and last lesson of roughly one month.

Table 6.8. Test no. 2: schedule & plan

<table>
<thead>
<tr>
<th>Date</th>
<th>Classes</th>
<th>Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>29/11/2016</td>
<td>1F</td>
<td>1N</td>
</tr>
<tr>
<td>06/12/2016</td>
<td>1F</td>
<td>1N</td>
</tr>
<tr>
<td>13/12/2016</td>
<td>1F</td>
<td>1N</td>
</tr>
<tr>
<td>10/01/2017</td>
<td>1F</td>
<td>1N</td>
</tr>
</tbody>
</table>

The environment in which we had the 4 lessons was the school library, which was quite big and dispersive, students did not have desks but just chairs. We organised the environment in small spaces, putting chairs in semi-circle. We had at our disposal a brand new interactive multimedia whiteboard. Some pictures of the environment and of the technical preparation of the room are shown in Figure 6.4 and 6.5.
6.3 Ecological test run no. 2

Figure 6.4. Library set-up

(a) Access point

(b) Switch

(c) Multimedia Whiteboard

(d) Tablets

Figure 6.5. The environment set-up

In the following we describe how we organised the four lessons.

Lesson 1

During the first lesson the teacher introduced the game to the class for roughly 20 minutes briefly explaining roles, tasks and objectives. Then, a test on the pre
6.3 Ecological test run no. 2

competencies and knowledge was done by the teacher. The test has been prepared by the teacher.

In order to have, from the first lesson, a gamified approach to the entire experiment, pre test was conducted in an unconventional way. In fact, Kahoot! was used to perform the test. Kahoot! is a free game-based learning platform, played by over 50 million people in 180 countries. It is designed for social learning, with learners gathered around a common screen. When a Kahoot! session is launched, it has its own unique PIN randomly generated by the system. Using any device with a web browser, players visit the website dedicated for quiz participation, kahoot.it, and input the unique PIN, and then enter a nickname before being directed to the waiting screen (the "lobby"). When all players’ nicknames appear in the lobby, the Kahoot! leader (normally the teacher) clicks the Start button to begin the quiz. During gameplay, the questions and up to 4 multiple choice answers are displayed on the main screen. Every answer corresponds to a distinctive color and shape. There are at most four rectangles on the screen of the learners’ devices, with the color and the shape on each, and the learner needs to click or tap on the rectangle representing the correct answer. The game design is such that the players are required to frequently look up from their devices, enabling social interaction with the teacher and their peers. For every question answered correctly, a number of points is awarded to each learner. At the end of every question, the learner sees his current score and his ranking in the class on his own screen. The learners with the 5 highest scores are displayed on the leader board on the main screen as well, to encourage healthy competition among players.

![Figure 6.6. Kahoot! pre test](image_url)

After concluding the Kahoot! session, students were already very excited and in a good mood, despite the results were not so exciting.

The teacher, after commenting the results of Kahoot! with the class, asked students to form teams and choose roles. Students get used to the idea of roles very fast and role were assigned in few minutes. We decided that students should possibly play in pairs, so three teams were formed, with 7 students each.

Due to lesson time constraint, teams and students registration was done by us so that at the beginning of the following lesson everyone was ready to play.
Lesson 2

During the second lesson students were divided per team and the teacher reminded the students about the few rules of the game. Teacher started the game session from PA firstly and then launched the PVB. Students logged in the 3DVG, everything works properly and players were able to play (Figure 6.7(a)) for roughly 40 minutes. The teacher was able to follow the game using the PVB, as shown in Figure 6.7(b).

(a) Students play the game  (b) Teacher follow the game on PVB

Figure 6.7. Pictures of lesson 2

Lesson 3

In lesson, we adopted the so called jigsaw technique. This is a method of organizing classroom activity that makes students dependent on each other to succeed. It breaks classes into groups and breaks assignments into pieces that each group assembles to complete the (jigsaw) puzzle. It was designed by the social psychologist Elliot Aronson [4] to help weaken racial cliques in forcibly integrated schools. The technique suggests to split classes into mixed groups to work on small problems that the group collates into a final outcome. We applied the jigsaw technique as follows. During this lesson students were not divided per team but per role, in fact we asked the teacher to organise groups of physicists, chemists, geologists and astro-biologist. Students were allowed to discuss problems and help each other, even if not belonging to the same team. In this configuration we wanted to encourage collaboration and knowledge sharing across teams.

The teacher reminded the students about the few rules of the game and players just started playing the game. We noticed that players started using some functionalities of the game just in this lesson. For example, the use of the IPA and the use of Knowledge Forum were not explored during the first game sessions. During this lesson Opedia® was used as well, Figure 6.8.
Lesson 4

During the fourth lesson students had 30 minutes to conclude the game and arrive to a final decision, with their team. Unfortunately, we discovered some bugs so some roles were not able to send notes and to upload the final tasks’ solution. However roughly half of the players were able to accomplish the assigned individual tasks. Due to time constraints students have not been able to discuss all together the team task so they just hazarded a possible solution. We realised that at least another lesson was needed but we proceeded with post test in order to assess the summative outcomes.

Teacher, using the PA interface, was able to look into details at the score to understand what players achieved during the game (Figure 6.9). From the score it was possible to infer how many questions players asked to the IPA, how many notes they added, how many quizzes they solved, etc.

![Figure 6.9. Screen-shot of players score seen from PA.](image.png)
6.3 Ecological test run no. 2

6.3.2 Learning assessment

At the beginning of this chapter we mentioned that students’ assessment in SGs can be of two types, completion assessment and in-process assessment. The first one corresponds to summative, while the second to formative assessment. Data about students achievements have been collected in two ways. First we monitored the progress in the game (formative) and at the end of the learning path we tested the overall learning outcomes (summative) via post test.

Formative Assessment

We monitored the formative outcomes of the single lesson tracing the game score. The score is divided into single items which give an indication of student’s learning (cfr. Table 4.1):

- Quiz
- IPA
- Tasks
- Notes

In the first two sessions of the games students were able to answer all quizzes. Quizzes content were preparatory for the task topics. Analysing the score, the teacher can see if the students answered correctly at the first, second or third attempt. The teacher cannot see which questions have been asked to the IPA but just how many. This value gives an indication of how much content students tried to investigate. Notes, validated by teacher, were also a good mean to test students achievement during the game. In fact, students shared their ideas and discoveries via notes and teacher was able to evaluate if the note was in line with the knowledge acquisition.

Figure 6.10, shows three screen-shots of the PA panel taken during the three game sessions (Lesson 2, 3 and 4), for the same student. We can see that students answered all quizzes in the first lesson. Quizzes score goes from a minimum of 0 (no quizzes answered) to a maximum of 15 (all quizzes answered correctly at first attempt). Analysing the score we can see that the student, in the second lesson, asked several questions to the IPA, in fact players gain 5 points for each question asked. In the final lesson the student got a note approved, which means that he discovered a piece of knowledge and shared it with his team-mate.
6.3 Ecological test run no. 2

![Score of team the mask 2nd lesson](image1)

![Score of team the mask 3rd lesson](image2)

![Score of team the mask final lesson](image3)

Figure 6.10. Example of formative assessment during the experiment

Unfortunately, in the 3DVG the task panel for uploading the final solutions was not working properly, so students that found the correct solutions to the assigned tasks reported verbally to the teacher and the data were not recorded into the score system. However, we annotate that roughly half of the students were able to solve the assigned tasks.

We verified that students followed a specific learning path, which was:

1. Explore 3D environment
2. Collect objects
3. Answer quizzes and gain access to simulation rooms
4. Enter simulation rooms and read assigned task
5. Ask help (IPA, OPedia)
6. Try to solve tasks
7. Share notes
8. Upload tasks solutions

Points 2 and 3 were interchangeable but were mandatory (by game mechanics) to, respectively, access tasks panels and enter simulation rooms. Formative assessment was quite satisfying because the majority of students followed the above steps and arrived to the final solution.
6.3 Ecological test run no. 2

Summative Assessment: pre and post test

The summative evaluation has been done via a post competence test and via a questionnaire of auto-assessment. With this experiment we didn’t want to prove that SGs are better than formal learning but that, proven the equality of outcomes, they instil the same curiosity in learners and enhance some of the so called 21\textsuperscript{st} century skills. As a consequence of that hypothesis, we tried to measure not just the competences per se, but the side skills as well.

We report hereafter the questions we asked and the percentage of correct answers pre and post test.

<table>
<thead>
<tr>
<th>Question asked</th>
<th>Skill</th>
<th>Pre test % correct answers</th>
<th>Post test % correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why we talk about Integrated Science?</td>
<td>Interdisciplinary</td>
<td>17,5</td>
<td>62,5</td>
</tr>
<tr>
<td>Which scientists should be part of a space mission that study exo-planets?</td>
<td>Interdisciplinary</td>
<td>20,0</td>
<td>45,0</td>
</tr>
<tr>
<td>Who is an astrobiologist?</td>
<td>Curiosity</td>
<td>63,5</td>
<td>44,0</td>
</tr>
<tr>
<td>What is an exoplanet?</td>
<td>Curiosity</td>
<td>71,0</td>
<td>40,0</td>
</tr>
</tbody>
</table>

From Table 6.9 we can see how there was an increment of the percentage of correct answers in interdisciplinary issues. However, we realised that the questions targeted at understanding the side skills were too few and vague. For this reason, in the third run we elaborated - together with two pedagogues - a more targeted set of questions.
Table 6.10. Summative assessment of competences: questions’ examples

<table>
<thead>
<tr>
<th>Question asked</th>
<th>Topic</th>
<th>Pre test</th>
<th>Post test</th>
</tr>
</thead>
<tbody>
<tr>
<td>How can we calculate gravitation?</td>
<td>Physics</td>
<td>33.5</td>
<td>50.0</td>
</tr>
<tr>
<td>After how many hours men die of hypothermia?</td>
<td>Astro-biology</td>
<td>24.3</td>
<td>45.0</td>
</tr>
<tr>
<td>Why planet magnetic field is important?</td>
<td>Geology</td>
<td>50.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Which are the three major gases of the terrestrial atmosphere?</td>
<td>Chemistry</td>
<td>0.0</td>
<td>30.0</td>
</tr>
</tbody>
</table>

An example of the questions we asked is reported in Table 6.10. The overall results of pre and post test are reported in Figure 6.11.

![Percentage of correct answers](image)

**Figure 6.11.** Summative results of the second test run

The results of the summative assessment suggest that students feebly learned. However, those results are fully compliant with the overall trend of the two classes. In fact, unfortunately, just 30% of the students reach sufficiency in the involved topics of study. Which means that, after a regular frontal lesson, a post test gives very poor results. Another important aspect that emerged during the lessons was that students had a very poor basic preparation. Game’s contents have been prepared having in mind the specific program of the target school grade. However, we found that students were not aware of the majority of the tasks proposed in NE. That was confirmed by teachers and by questionnaire results, as explained next.
Questionnaires

In order to evaluate the effectiveness of NE and the perception of learning outcomes by students, we asked some questions via a questionnaire. We used a Likert scale, in which the format of the five-level Likert items ranged from *Yes a lot* to *Not at all*.

**Table 6.11.** Questionnaire results from test run no. 2: liking & learning

<table>
<thead>
<tr>
<th>Qn</th>
<th>Question</th>
<th>Yes, a lot</th>
<th>Yes, enough</th>
<th>Indiff.</th>
<th>Not much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>Did you enjoy playing the game?</td>
<td>2</td>
<td>19</td>
<td>7</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Q2</td>
<td>Do you think that playing the game is more engaging than a traditional lesson?</td>
<td>11</td>
<td>16</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Q3</td>
<td>Would you recommend to students from other classes to try the game?</td>
<td>7</td>
<td>11</td>
<td>9</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Q4</td>
<td>Do you think you have learned by playing the game?</td>
<td>1</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Q5</td>
<td>Do you think that playing the game has allowed you to learn in a more interesting way?</td>
<td>5</td>
<td>17</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Q6</td>
<td>Do you think that playing the game has allowed you to learn faster?</td>
<td>1</td>
<td>10</td>
<td>12</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

The first three questions we asked were very simple and aimed at establish if students liked the game and would recommended it to other students. Answers are displayed in Figure 6.12. 64% of students enjoyed playing the game while 15% disliked it. 81% of the students think that NE is more engaging then a traditional lesson but just 54% would recommend the game to other students while 27% are indifferent.
First, we asked the students if they learned by playing the game, just 36% replied in a positive way. We then asked if NE is better than a frontal lesson and if they think they learned in a playful way. 67% of the students prefer this type of teaching but just 33% think to have learned in a pleasant way.

We believe that the results are due to several reasons. We received a lot of positive feedback about the game from students, however they complain about English (the game was fully in English.) because they were not able to understand quizzes and tasks. For the same reason they did not ask many questions to the IPA. That was the main reason why we decided to rewrite all the quizzes and tasks in Italian and maintain just the IPA in English in the final version of NE.
6.3 Ecological test run no. 2

English level of Italian students is very poor. In fact, 38% of the students were discouraged by English, as reported in Figure 6.14.

![Figure 6.14. Language obstacle](image)

Secondly, another important remark coming from students was that there were many bugs in the 3DVG and sometimes they wasted time just having fun, remained trapped in a texture, or just jumping in places with no exit. In the final version of the 3DVG we spent a lot of time fixing malfunctions and bugs.

Finally, we found that roughly 50% of the students considered the assigned tasks too difficult and 41% considered the quizzes too difficult as well, see Figure 6.15.

![Figure 6.15. NewEarth contents adequacy](image)

Finally, we asked other questions aimed at evaluating single components of the game, as reported in Table 6.12.
6.4 Ecological test run no. 3

Table 6.12. Questionnaire results from test run no. 2: usability

<table>
<thead>
<tr>
<th>Qn</th>
<th>Question</th>
<th>Yes, a lot</th>
<th>Yes, enough</th>
<th>Indiff.</th>
<th>Not much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q7</td>
<td><em>Interfaces were clear enough?</em></td>
<td>1</td>
<td>16</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Q8</td>
<td><em>Do you think the PVB is useful?</em></td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Q9</td>
<td><em>Do you like seeing teams position on the PVB?</em></td>
<td>10</td>
<td>6</td>
<td>11</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Q10</td>
<td><em>Do you like seeing scores on PVB?</em></td>
<td>11</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Q11</td>
<td><em>Did you use Opedia?</em></td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Q12</td>
<td><em>Were Opedia insights useful?</em></td>
<td>4</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Q13</td>
<td><em>Was IPA useful?</em></td>
<td>0</td>
<td>10</td>
<td>9</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Q14</td>
<td><em>were adding and sharing personal notes useful?</em></td>
<td>7</td>
<td>8</td>
<td>19</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Q15</td>
<td><em>Did you like teacher to be responsible for note validation?</em></td>
<td>2</td>
<td>4</td>
<td>17</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Q16</td>
<td><em>Would you like to have a personal customizable avatar?</em></td>
<td>14</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

We used data in Table 6.12 to improve the 3DVG interface and Opedia contents, before the third run. The IPA remained in English and we did not add a personalised avatar. The latter two points will be discussed in Chapter 7.

Finally we asked two other questions, reported in Table 6.13.

Table 6.13. Questionnaire results from test run no. 2: video-game

<table>
<thead>
<tr>
<th></th>
<th>Yes, very much</th>
<th>Not much</th>
<th>Not at all (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Do you like video-game?</em></td>
<td>(23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Do you think that NE can be considered a video-game?</em></td>
<td>(13)</td>
<td></td>
<td>(14)</td>
</tr>
</tbody>
</table>

We did not find any correlation between the two answers. In this version we were not able to implement the unsolicited hints mechanism so we couldn’t test it.
6.4 Ecological test run no. 3

6.4 Ecological test run no.3 : Improved version of *NE*

In this section, we describe the ecological tests aimed at assessing the final version of *NewEarth*. The main changes in the 3DVG were:

- Quizzes and tasks in Italian
- Reduced number of objects to be collected (6 per player)
- Reduced point gained for IPA, notes and chat
- Unsolicited hints.

6.4.1 Ecological environment description

In this test run, we worked with two classes formed by 24 and 29 students, as reported in Table 6.14. The two classes were in the same school but they differs in the specialization with respect to the previous run.

<table>
<thead>
<tr>
<th>School Name</th>
<th>Classes</th>
<th>Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITI Bernini</td>
<td>1H (Scientific Institute)</td>
<td>29 (9♀, 20♂)</td>
</tr>
<tr>
<td>ITI Bernini</td>
<td>1D (Industrial and Technical Institute)</td>
<td>24 (7♀, 14♂)</td>
</tr>
</tbody>
</table>

In Table 6.15 we report lessons schedule and plan. We had 5 lessons for each class between February and March 2017.

<table>
<thead>
<tr>
<th>Date</th>
<th>Classes</th>
<th>Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>13/02/2017</td>
<td>1H</td>
<td>1D</td>
</tr>
<tr>
<td>20/02/2017</td>
<td>1H</td>
<td>1D</td>
</tr>
<tr>
<td>27/02/2017</td>
<td>1H</td>
<td>1D</td>
</tr>
<tr>
<td>06/03/2017</td>
<td>1H</td>
<td></td>
</tr>
<tr>
<td>13/03/2017</td>
<td>1D</td>
<td></td>
</tr>
<tr>
<td>13/03/2017</td>
<td>1D</td>
<td></td>
</tr>
<tr>
<td>17/03/2017</td>
<td>1H</td>
<td></td>
</tr>
</tbody>
</table>

In this test run, we collaborated with two researchers from the Pedagogy department, specialised in new technologies and learning. One person followed all the lessons taking notes about the learning aspect of the game and about social dynamics among students, related to the game. They prepared a specific questionnaire that
helped us in investigating one of the main objectives in the pedagogical area of this thesis: the acquisition of side skills.

The environment in which we had the 4 lessons was a classroom so called 3.0 with special refurbishment. Beside the multimedia whiteboard, special desks and chairs were available as pictured in Figure 6.16.

![Figure 6.16 Pictures of environment set-ups](image)

The special refurbishment shown in Figure 6.16, has been the outcome of an agreement with the private company Eugeni Tecnologie srl. They provided tablets, multimedia whiteboard (2nd and 3rd) and refurbishment for the 3rd run. Eugeni is also the owner of Opedia® and gave us the licences to use it for the experiments.

That special furniture gives the possibility, to the teacher, to easily work in groups and change class configuration when needed. The desks have a trapezoidal form, can be aligned or put in circle as shown in Figure 6.16(d). That configuration is the one we used for this test run.

**Lesson 1**

During the first lesson, the teacher introduced the game to the class for roughly 20 minutes briefly explaining roles, tasks and objectives. Then, a test on the pre competencies and knowledge was done by the teacher. The test was the same as in test run no. 2 and the teacher used Kahoot! as well to evaluate pre competencies.

Students were allowed to choose places and they split in teams. Figure 6.17 gives an idea of an ideal game session in class. In this test run, we formed 4 teams instead of 3. We divided the two classes, named 1D and 1H, respectively in teams of 6 and 7 players. The idea was, as in the previous run, to play in couple when possible.
Lesson 2
In this test run we did not create accounts in advance, but we asked the teacher and the students to do it, following this workflow.

- Students had 5 minutes to decide the name of the team.
- Once they decided, the teacher added teams using the PVB, as shown in Figure 6.18.
- Students registered and hooked to a specific already created team. During the registration students chose their roles.

The workflow previously described worked fine and in 15 minutes all students were registered to a team, see Figure 6.19(a). The remaining 30 minutes students played the game, see Figure 6.19(b).
6.4 Ecological test run no. 3

Lesson 3

In this lesson we made a small change to configuration of the desk. In fact, we had the possibility to change one desk per group with a coloured one. We had four different coloured desks: blue, yellow, orange, green. We decided to use colours and we told students that the group’s master should seat at the coloured desk. In this lesson, we divided students per role and not per team, following the already described jigsaw technique. We had a group of tables for geologists, one for physicists, one for astro-biologists and one for chemists. Again, we wanted to encourage cross collaboration among teams and have discussions about the individual tasks, common to all roles.

We noticed that students used IPA and Knowledge Forum very rarely. Once they understood that asking questions and adding notes made them earn points, they started to use them more.

In this lesson, students played roughly 45 minutes.

Lesson 4

In this lesson students were asked to accomplish the assigned individual and team tasks. They were divided per team and they nominated a master that was responsible for collecting data from players, coordinate the discussion and compile the final
6.4 Ecological test run no. 3

report. The master seated in the coloured desk and he/she was required to take timing and make decisions.

![Image](image1.png) ![Image](image2.png)

(a) Team’s master coordinates activities  (b) Student finish his task

**Figure 6.21.** Pictures of lesson number 4

**Lesson 5**

Questionnaires were compiled by students of both classes, but in different ways. With students from 1D, we:

- Discuss achievements and results
- Compile questionnaires
- Test post competences

In order to test post competencies we again decided to use Kahoot! to maintain the gamified approach. Picture of students, that clearly shows engagement, are reported in Figure 6.22.

![Image](image3.png) ![Image](image4.png)

(a)  (b)

**Figure 6.22.** Students playing Kahoot!

Class 1H compiled the questionnaire as part of a standard lesson. We did not use Kahoot! but a standard written test. Questions in both classes were the same.

**6.4.2 Learning Assessment**

Data about students achievements have been collected in two ways. First we monitored the progress in the game (formative) and at the end of the learning path we tested the overall learning outcomes (summative) via post test.
6.4 Ecological test run no. 3

Formative Assessment

We monitored the formative outcomes of the single lesson tracing the game score as in the previous run.

The schema followed by students during the game sessions is the same described in the previous run. In this run, students were also able to upload the task solutions, so we were able to monitor the final achievement, as reported in Figure 6.10. In PA screen-shots we can see how many students answered the final tasks (marked with a red circle).

![Example of task score, team witch](image1)

(a) Example of task score, team *witch*

![Example of task score, team hard rock](image2)

(b) Example of task score, team *hard rock*

**Figure 6.23.** Example of formative assessment during run no.3

The maximum score for tasks is 30 (10 per each task). Point earning differs if the students replied correctly at the first attempt. We know that, for example in Figure 6.23(b), students gave two correct answers at the first attempt, while in Figure 6.23(a) students gave just one correct answer at the second attempt.

**Table 6.16.** Overall achievement per team

<table>
<thead>
<tr>
<th>Class</th>
<th>Team name</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>The magic</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Ultras</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Hard Rock</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Witch</td>
<td>43</td>
</tr>
<tr>
<td>1H</td>
<td>Dream Team</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>Rosa Fashion</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>Dpg777</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Gnob</td>
<td>60</td>
</tr>
</tbody>
</table>

The results presented in table 6.16 correspond to the average score achieved by each team. The maximum individual score that was possible to reach in the game was 75. The results reveal that almost every team was able to follow the game,
however the following table gave us a better understanding of the score. In fact, we were able to trace score for each game’s attitude.

Table 6.17. Detailed percentage of score

<table>
<thead>
<tr>
<th>Class</th>
<th>Quiz</th>
<th>Task</th>
<th>Objects</th>
<th>Notes</th>
<th>Ipa</th>
<th>Chat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>85</td>
<td>35</td>
<td>90</td>
<td>20</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>1H</td>
<td>90</td>
<td>40</td>
<td>92</td>
<td>25</td>
<td>55</td>
<td>8</td>
</tr>
</tbody>
</table>

The percentage reported in table 6.17 gives a clear indication of what was easy/difficult in the game. The easy part for the majority of the students was to answer quizzes and to find 3D objects around the spaceship. Just a minority of them (37%) was able to solve the assigned tasks. Very few students used notes and chat; roughly 50% asked questions to the IPA. Two major concerns were expressed by students. First a technical difficulty, due to game not properly designed, was verbally reported to the teacher at the end of the last lesson: students were not able to insert the final solutions they found. In fact, calculus were done on paper (as depicted in Figure 6.21) and then should be uploaded to a final panel that was tricky to discover. Second, students were not able to formulate questions in English to the VT and that justifies the percentage of low questions asked. The majority of the students declared they were keen on asking questions in Italian.

Summative Assessment: pre and post test

The summative evaluation has been done via a post competences test and via a questionnaire of auto-assessment.

As mentioned, questions asked were exactly the same of run no.2 so hereafter we report just the overall obtained results.

Unfortunately, results are not so good. We asked the teacher of the two classes to help us interpret the data and she was not surprised at all. In fact, she explained that more then half of the students (precisely 16 in 1H and 14 in 1D) have insufficient
marks, as in the second run. In addition, during the pre test, we noticed that the level of knowledge was really low. We expected the game to better understand the topics that were already explained and not to completely replace a frontal lesson.

Questionnaires

We used the same questionnaire of test run no. 2. Table 6.18 reports the overall results.

Table 6.18. Questionnaire results from test run no. 3

<table>
<thead>
<tr>
<th>Qn</th>
<th>Question</th>
<th>Yes, a lot</th>
<th>Yes, enough</th>
<th>Indiff.</th>
<th>Not much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td><em>Did you enjoy playing the game?</em></td>
<td>7</td>
<td>23</td>
<td>5</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Q2</td>
<td><em>Do you think that playing the game is more engaging than a traditional lesson?</em></td>
<td>16</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Q3</td>
<td><em>Would you recommend to students from other classes to try the game?</em></td>
<td>8</td>
<td>21</td>
<td>4</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Q4</td>
<td><em>Do you think you have learned by playing the game?</em></td>
<td>4</td>
<td>26</td>
<td>0</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Q5</td>
<td><em>Do you think that playing the game has allowed you to learn in a more interesting way?</em></td>
<td>4</td>
<td>26</td>
<td>0</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Q6</td>
<td><em>Do you think that playing the game has allowed you to learn faster?</em></td>
<td>13</td>
<td>19</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>
We found that the percentage of likeability (Table 6.25) is more or less the same as the second run. However, we found that learnability (Table 6.26) perception has improved a lot. In fact, we passed from 37% to 64%. That is probably due to the Italian version and improved quality of game and contents. Unfortunately this was not confirmed by the summative assessment, as it will be further deepened in Chapter 7.
Table 6.19. Questionnaire results from test run no. 3: usability

<table>
<thead>
<tr>
<th>Qn</th>
<th>Question</th>
<th>Yes, a lot</th>
<th>Yes, enough</th>
<th>Indiff.</th>
<th>Not much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q7</td>
<td>Interfaces were clear enough?</td>
<td>6</td>
<td>23</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Q8</td>
<td>Do you think the PVB is useful?</td>
<td>12</td>
<td>18</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Q9</td>
<td>Do you like seeing teams position on the PVB?</td>
<td>12</td>
<td>16</td>
<td>12</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Q10</td>
<td>Do you like seeing scores on PVB?</td>
<td>12</td>
<td>16</td>
<td>12</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Q11</td>
<td>Did you use Opedia?</td>
<td>6</td>
<td>14</td>
<td>8</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>Q12</td>
<td>Were Opedia insights useful?</td>
<td>4</td>
<td>11</td>
<td>16</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Q13</td>
<td>Was IPA useful?</td>
<td>4</td>
<td>13</td>
<td>6</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Q14</td>
<td>Were adding and sharing personal notes useful?</td>
<td>10</td>
<td>15</td>
<td>12</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Q15</td>
<td>Do you like teacher was responsible for note validation?</td>
<td>8</td>
<td>13</td>
<td>10</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Q16</td>
<td>Would you like to have a personal customizable avatar?</td>
<td>26</td>
<td>12</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

We report in Table 6.20 a comparison of data from the two test runs. In order to quickly understand the trend we summed the results for [Yes, a lot - Yes, enough] and [Not much - Not at all], and labelled respectively Positive and Negative.

Table 6.20. Comparison between run no. 2 and run no. 3

<table>
<thead>
<tr>
<th>Qn</th>
<th>Positive</th>
<th>Neutral</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Run 3</td>
<td>Run 2</td>
<td>Run 3</td>
</tr>
<tr>
<td>Q7</td>
<td>62↑</td>
<td>53</td>
<td>13</td>
</tr>
<tr>
<td>Q8</td>
<td>64↑</td>
<td>47</td>
<td>13</td>
</tr>
<tr>
<td>Q9</td>
<td>60↑</td>
<td>50</td>
<td>26</td>
</tr>
<tr>
<td>Q10</td>
<td>60↓</td>
<td>72</td>
<td>26</td>
</tr>
<tr>
<td>Q11</td>
<td>36↓</td>
<td>47</td>
<td>17</td>
</tr>
<tr>
<td>Q12</td>
<td>32→</td>
<td>38</td>
<td>34</td>
</tr>
<tr>
<td>Q13</td>
<td>36→</td>
<td>31</td>
<td>9</td>
</tr>
<tr>
<td>Q14</td>
<td>53→</td>
<td>47</td>
<td>26</td>
</tr>
<tr>
<td>Q15</td>
<td>45↑</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>Q16</td>
<td>81↑</td>
<td>69</td>
<td>11</td>
</tr>
</tbody>
</table>
Students found the interface more clear and that is due to the effort we have made in correcting errors in the 3DVG. We believe students of the third run used the PVB better, interacting more with the teacher. That was probably due also to the environment more suitable for the lessons. The pick in question Q10 in run no. 2 is due to the fact that we implemented this feature just for the last lesson and students were very happy about this novelty. Opedia was not used very much. Also if the content is targeted to the game, students preferred using Internet, that was allowed as well. IPA was under used because of the language, in fact many students reported they didn’t feel confident in asking questions in English. The strong majority of students wished to have a personal avatar to customise in the game.

Finally we asked two other questions, reported in Table 6.21.

Table 6.21. Questionnaire results from test run no. 2: video-game

<table>
<thead>
<tr>
<th>Do you like video-games?</th>
<th>Yes, very much (33)</th>
<th>Not much (10)</th>
<th>Not at all (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you think that NE can be considered a video-game?</td>
<td>Yes, very much (20)</td>
<td>Not much (7)</td>
<td>I don’t know (20)</td>
</tr>
</tbody>
</table>

Again, we did not find any correlation between the two answers.

As anticipated at the beginning of this section, we worked with two pedagogues and, beside the annotations during game sessions they prepared a questionnaire aimed also at verifying the acquisition of side skills. The questionnaire used a Likert scale and has been dispensed to both classes. Questionnaire outcomes are reported in Table 6.22. All statements started with "I learned to:"
### 6.4 Ecological test run no. 3

Table 6.22. Pedagogical questionnaire results from test run no. 3

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes, very much</th>
<th>Yes, a lot</th>
<th>Enough</th>
<th>Not much</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>break-down tasks into small problems</td>
<td>3</td>
<td>9</td>
<td>16</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>appreciate collaborative work</td>
<td>6</td>
<td>15</td>
<td>10</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>appreciate technology potential</td>
<td>2</td>
<td>10</td>
<td>13</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>view topics from a different perspective</td>
<td>1</td>
<td>12</td>
<td>14</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>achieve complex tasks in a collaborative way</td>
<td>5</td>
<td>13</td>
<td>15</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>plan collaborative work</td>
<td>5</td>
<td>10</td>
<td>21</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>understand requirements coming from different people</td>
<td>7</td>
<td>8</td>
<td>14</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>develop an idea in a collaborative way</td>
<td>4</td>
<td>13</td>
<td>15</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>understand the importance of work in progress</td>
<td>1</td>
<td>10</td>
<td>22</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>understand the importance of interdisciplinary</td>
<td>6</td>
<td>16</td>
<td>15</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

The results reported above can be better understood if expressed in percentage, see Figure 6.27.

![Side skills evaluation](image)

Figure 6.27. Side skills evaluation
6.4 Ecological test run no. 3

Positive answers are above 60%, in some cases arriving at roughly 70%. In particular, collaboration, interdisciplinary and problem solving were felt by students as positive outcomes of the game.

6.5 Unsolicited hints

The hints mechanism has been just partially tested. In order to evaluate it, instead of asking players if they found the hints useful, we gave them the opportunity to vote thumb up or down (see Figure 5.23). The system registered all the variables’ values, the provided hint and the players’ vote.

In Figure 6.28 we reported an example of log file.
6.4 Ecological test run no. 3

We study acceleration in cinematics while we study force in dynamics. Newton's universal gravity law describes how the force is related to the planet's masses.

If you have a pendulum you can derive the gravity acceleration with a chronometer, a meter and knowing the law that link the period of swing, the length and the gravity acceleration.

If you have a pendulum you can derive the gravity acceleration with a chronometer, a meter and knowing the law that link the period of swing, the length and the gravity acceleration.

Just two elements forms the 99% of the terrestrial atmosphere!

When the sunlight hits the atmosphere many things happen, depending on the wavelength of the light. Each wavelength behaves in different ways: some are reflected, some absorbed and some others reach the Earth.

All four of the innermost, rocky planets of the Solar System may have had large variations of their obliquity in the past.

It has been shown that the obliquities of exoplanets in the habitable zone around low-mass stars tend to be eroded in less than a billion years, which means that they would not have seasons as Earth has.

Earth’s obliquity may have been reasonably accurately measured as early as 1100 BC in India and China. The ancient Greeks had good measurements of the obliquity since about 350 BC. About 830 AD, the Caliph Al-Mamun of Baghdad directed his astronomers to measure the obliquity, and the result was used in the Arab world for many years.

Figure 6.28. Log files of unsolicited hints record

In the figure above, it is possible to evaluate user's assessment of hints: we coloured in green true vote and in red the false ones. True means thumb up and false thumb down. We have underlined, in yellow, records that show emotional detections. In one game session, for all users (14 because they played in pairs when possible), we provided 72 hints, 47 were judged positive and 25 negative. An interesting result that we found, even if validated on this small sample is that when players’ emotion
was negative they always replied thumb up to the provided hint, when emotion was positive they always replied thumb down to the provided hint. Table, summarised found results.

**Table 6.23.** Test no. 3: unsolicited hints assessment

<table>
<thead>
<tr>
<th>Total hints</th>
<th>Thumb up</th>
<th>Thumb down</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>65%</td>
<td>35%</td>
</tr>
</tbody>
</table>

**Detected player’s emotion**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

To have a better understanding of unsolicited hints provisioning we have to make the following annotations:

- All the provided hints were in English and students complained about that
- The hints were not linked to the non-explored contents but were provided randomly
- Emotion detection has been limited because of the poor usage of IPA and chat.

From this preliminary evaluation, we have an indication that emotions’ detection can be a valuable mechanism to provide help during game sessions.
Chapter 7

Conclusions

In this chapter we report some considerations about the work done so far and future works. We must approach everything as a child does, assuming nothing, observing everything, and listening with a waiting open soul. “Every truth has four corners: as a teacher I give you one corner, and it is for you to find the other three.” Confucius

In this concluding chapter we summarise our work and contributions; we introduce some lessons learned to further discuss possible future work and research directions.

7.1 New vision in education

In this research we have outlined the importance of a new vision in education. The vision should foster both technical and pedagogical issues. From a technical perspective, exploiting AI in education it is possible to implement Intelligent Instructional Systems and modelling them in several ways. We have chosen to follow the Game Based Learning trend and to implement a Serious Game to teach STEM. The SG has been developed as a Virtual Learning Environment in which an Intelligent Pedagogical Agent follows the learners in their learning path. Emotions enter the loop because it is important to construct a trusted relationship between students and IPAs, which helps filling the lack of the human presence. However, our SG has been thought to be used in class, as a supporting teaching tool. Thus, we imagined an interface both for students and teachers, the latter being fully part of the educational process. Technology Enhanced Learning should help to integrate 21st Century Skills in the new learning process, which are required to students to afford nowadays challenges. We built NewEarth, a SG that is also a role-playing game, in which students are part of a team of scientists that should find another planet where humans can live. NE does not want to substitute the traditional
7.1 New vision in education

lessons, nor claims to teach in a more proficient way. *NE* is an integration to the traditional lessons, offering a new tool that supports the learning process throughout the necessary teaching time. In fact, *NE* is not a simulation tool or a game to be used una tantum, it aims at assisting learners and teachers in a process that usually lasts several lessons. The optimal assessment should last at least the same period that teacher usually teaches a specific topic. Even if we cited Michael and Chen statement that "Serious Games, like every other tool of education, must be able to show that the necessary learning has occurred", some important authors tend to disagree with this statement for some reasons. In fact, there have been many so called tools of education that were unable to demonstrate significant learning (see, e.g.: The No Significant Difference Phenomenon \(^1\)). Nevertheless, these tools continue to be produced and used successfully even without convincing comparison with other learning methods, at least because most often they do not replace but complement them. This observation is particularly true in the case where "other methods" are simply not available, as it is the case in "isolated" situations: lack of teachers, lack of expertise, lack of time etc. An important variable influencing learning is not the teaching method (behaviourist vision), rather the learner’s motivation (constructivist vision). Additionally Serious Games are supposed to induce learning as a side effect of playing, a kind of informal learning. Learning goals are therefore not easily specified, consequently hard to measure. We adopted a stratified strategy to test our work. We first assess the validity of the VT in a stand alone application and then we integrate it in the Serious Game. In a second step, we measured the learning outcomes not only analysing the *NewEarth* results but also dispensing questionnaire pre and post games’ sessions. Our assessment (4/5 lessons in formal learning education) gave positive results and a clear indication that students appreciate a new way of providing lessons. In addition we had a clear indication that 21\(^{st}\) century skills can be addressed by our SG. In fact, most students indicate that they learned the importance of collaboration, problem solving and interdisciplinary issues.

From an engineering point of view, we put a lot of effort in integration-interoperability of several different available software tools and a few gluing algorithms. Among them: Ajax, REST-based web services, HTML-CSS, DOM, Javascript, XMLHttpRequest, Unity3d, C#, Stanford CoreNLP, CrazyTalk, IBM Bluemix Watson. As Peter Wegner said in \([105]\) "The paradigm shift from algorithms to interaction captures the technology shift from mainframes to workstations and networks, from number-crunching to embedded systems and graphical user interfaces, and from procedure-oriented to object-based and distributed programming. The radical notion that interactive systems are more powerful problem-solving engines than algorithms is the basis for a new paradigm for computing technology built around the unifying concept of interaction." In order to simplify the interaction among several components of *NewEarth* we decided to develop the VT as a service. That gave us the possibility to have the VT in the server side and consequently to lighten the client side. In addition, in the future plan we would like the VT to support what is traditionally known as application customization. In other words, a single user (games’ viewer) can alter the set of configuration options (a.k.a. parameters) that affect its functionality and look-and-feel. Each student may have his/her own settings for the configuration options. The VT can be customized to the degree
it was designed for based on a set of predefined configuration options. Two other
important features of the game that deserve more attention in future developments
are the authoring tool, the knowledge forum and the unsolicited hints mechanism.
Both topics open to further investigation because, despite their potentialities, few
literature have been found. The authoring tool that we developed is a web applica-
tion where an author (a teacher or a content expert) can register and, once verified,
add his content. The goal is to make authoring intelligent dialogue managers simple
for teachers. That has been done in the past by Cerri in [16] and Woolf in [108].
For users not always confident with complex interface and contents’ organization
in hundreds of categories and rules, will be very helpful to provide their content
(because field experts) in a simplified way. The idea of the Knowledge Forum was
that each student could add a note and read the notes added by other students
once they arrived in the PVB. The intent of such mechanism was to encourage
the collaborative learning, like the jigsaw techniques that we have introduced in
one lesson during the second and third runs. Notes are stored just in the team’s
Knowledge Forum, so once read in the PVB, they can be later accessed and used
as a support to arrive to the final solution just by the team itself. However, this
mechanism has been criticized by students that did not want to share information
among teams because this curbs the competition, which is one of the important
lever of games. In the future, we have to revise this mechanism and probably keep
the notes just for the team’s members. The unsolicited hint mechanism has been
addressed just at a first stage. It will be very interesting to deepen the way in
which hints could be provided based on users’ behaviours during games’ sessions.
Our first attempt has been to catch emotions in a very simple way (analysing text
messages) but more powerful tools, such as sensors or more complex user modelling
(like Conati did in [21]), can be integrated as well.

7.2 Lessons learned

In this section, we introduce the lessons learned, which should be used to improve
future work. The lessons learned concern mainly the assessment process that should
be as much accurate as possible. In fact, in the case of formal learning environment, it
is very important to compare the results from two samples of students to understand
which are the added values of the proposed experiments. As stated in Chapter
6, assessment in education is very complex because learning is complex. In fact,
learning requires time to understand and time to elaborate new information. Beside
that, learning the so called side skills is maybe less complex if supported by good
teaching practice but however very difficult to assess, especially in the time frame of
an experiment.

The first lesson learned concerns how an experiment should be arranged.

First lesson

It is useful to have a control and experimental group in order to compare the
results obtained in both groups.

\(^1\)http://nosignificantdifference.org/
This is necessary to validate research questions and objectives in a more comprehensive way (Zakharov et al. in [113]).

The second lesson regards the loop of the formal environment in which we teach.

<table>
<thead>
<tr>
<th>Second lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiments should be as close as possible to a real classroom lesson.</td>
</tr>
</tbody>
</table>

It is very important to introduce a non-invasive assessment of knowledge and learning progress (Kickmeier-Rust & Albert in [53]).

<table>
<thead>
<tr>
<th>Third lesson</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is necessary to have two versions of the same system, one basic and one with novelties</td>
</tr>
</tbody>
</table>

In fact, it is not enough to test the system with control and experimental group. The system itself should be provided with basic functionalities and with novelties that could be evaluated per se. For example, we could have tested the Android app with a basic IPA and with and emotional IPA to better understand the importance of emotions in learning. The same is true for the unsolicited hints mechanism: we should have tested a version with randomly generated hints and another version with our algorithm. In that way the goodness of the system could be better evaluated. However, due to the complexity of our system and the related assessment we could not manage too many tests. The latter reasoning leads us to the future works.

### 7.3 Future works

The results obtained during the assessment process are encouraging. Students enjoyed NewEarth and they claimed it was more interesting than a frontal lesson, they would like to play more and they participated actively giving feedbacks and suggestions outside the questionnaire, just chatting with us at the end of each lesson.

The plot of the game has been appreciated and students enjoyed to be someone else, scientists with a specific challenge.

In the future, we have to work on some aspects that should be improved. The IPA should be able to chat in Italian, otherwise the majority of students will be discouraged a priori to ask questions.

The unsolicited hints mechanism seems very interesting. However we should probably change the monitored nodes because, for example, the chat was not used much. In addition, we should provide hints related to context that has not yet been explored. It means that, if a student already answered a quiz or solved a specific task, hints should be provided accordingly.

Another type of help that students required, was linked to what to do in the game. So, another variable that could be taken into account in the Bayesian network is, for example, the player’s position. Monitoring this variable as well, we can provide hints
7.3 Future works

helping them to find objects to be collected ("Turn left and then right to find the nylon string") or a panel to interact with ("Follow the light to discover the panel").

Game mechanics can be improved as well. Studying the collected data during real game sessions, we will be able to better integrate learning and fun. Especially, we would like to make the experiment more integrated into the 3D environment. For example, a very interesting suggestion that came from one of the students, was to move an avatar depending on the gravity of the planets.

Finally, another aspect that deserves attention is the possibility to have a customizable avatar. This request, that came from roughly 80% of the students, is perfectly in line with what claimed in the study by Dickey (see [28]), and reported in Table 3.1. In fact, one of the successful aspects of Massively Multiplayer Online Role-Playing Games is the possibility to have a customizable avatar that represents players in the game. However, in a Serious Game that should be played in class, time spent in gaming is very important for teachers and time dedicated to customizing an avatar is still an open question: is it serious enough?
Acknowledgement

I express my special appreciation and thanks to my advisor Professor Massimo Mecella, who has been a fundamental mentor for me. I thank him for encouraging my research and for allowing me to grow as a research scientist. His advices have been very important and helped me to not get lost in a very complex world of information and stimuli. I thank my co-advisor prof. Luigia Carlucci Aiello who entered me in the passionate world of Artificial Intelligence and who gave me precious advices and strong encouragement. I really benefited from her acute observations shared always with real participation.

I thank my husband and my three daughters (Melissa, Maya and Viola) from whom I borrowed the exo-planets’ names and was always inspired. My daughters are too young to appreciate all the efforts made to accomplish my research, but I hope they will be inspired, in their future, from my effort to keep them always involved.

I also thank all the students (from the university and high-school), teachers and principals I worked with that made the implementation and testing of NewEarth possible and enjoyable.

Finally, I would like to thank my parents who always supported my decisions, even if I am more than adult.
Appendices
Appendix A

Emotion & schooling project plan

In this section, we provide more details about the app design and implementation.

In Table A.1, we report the detailed description of the lessons that we had in the pilot project.

Table A.1. Emotion and schooling lessons’ work plan

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Teacher activities</th>
<th>Student activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>17/09/2014</td>
<td>During the first lesson we have presented the project to the students and we have explained some concepts related to games, virtual tutoring, animated agents and artificial intelligence. In this way, students were able to understand the context in which the project was placed. We did not yet mentioned anything about emotions.</td>
<td>Two students of the group were involved in a written test situation. At the end of the test students expressed their feeling and emotions. The students not involved in the test were committed to describe facial expression of their classmate.</td>
</tr>
<tr>
<td>02/10/2014</td>
<td>In this lesson, we presented to the students the form they should use throughout the project. We explained the meaning of each column of the form and we mainly explained and discussed basic emotions starting from the hereafter schema proposed by Lane in [57].</td>
<td>At the end of the lesson students were asked to do a written test without notice. The test was of medium difficulty and was about some topics explained few days before by the teacher. At the end of the test students filled the form we provided, each one with her own feelings. We left students with the freedom to express freely without taking into account the emotions already explained.</td>
</tr>
<tr>
<td>Date</td>
<td>Description</td>
<td>Details</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>16/10/2014</strong></td>
<td>In this lesson, we continued the work on emotion reading some literature texts and having free discussion among students. We also projected some video taken from Italian movies about everyday school life in which students and teachers were in exasperated and paradoxical situations in order to stimulate the students reasoning about the relationship with teachers.</td>
<td>Students filled another column of the provided form. For each one of the eight basic emotions identified in the previous lesson, students were asked to express the wished behaviour of a virtual tutor, namely the agent. Together with the students we limited the possible reaction of the agent to four: Encouragement, Suggestion, Joke, No intervention.</td>
</tr>
<tr>
<td><strong>30/10/2014</strong></td>
<td>In this fourth lesson, we discussed about emotions and corresponding facial expression. In order to make the topic funnier we used some materials coming from the famous tv series &quot;Lie to me&quot;. We invited a professional dabber to this lesson to work on pitch and facial mimics.</td>
<td>Students were invited to perform a psychology test in which they should recognise the emotion expressed in a picture. The test was online and interactive. Students were asked to answer the test all together favouring brainstorming and constructive discussion about emotions and facial mimics. In addition, they filled another column of the form in which they should invent a phrase that the agent should say in each specific student emotional state and based on the already expressed wished behaviour of the agent (encouragement, suggestion, etc). At the end of this lesson, each students has his form fully filled and ready to be used in the upcoming implementation phase of the project.</td>
</tr>
<tr>
<td><strong>28/11/2014</strong></td>
<td>The CrazyTalk tool has been presented and explained to the students.</td>
<td>Students freely used the already explained tool in order to get familiar with it. They tried out different functionalities of the tool like: agent creation, agent animation, registering phrases, etc.</td>
</tr>
</tbody>
</table>
Teachers asked the students to fill a very simple questionnaire about the agent they would like to have expressing some desiderata like: should be human or not, should resemble you or someone that you know, should be fantasy or really existed, etc. In this way, teachers collected results about the physical aspects of the agent. Results are reported below.

During the last two lessons students were grouped into group of two and started creating their personal agent using the form they already filled with emotions and phrases to animate the character they have chosen.

In Figure A.1, we have represented the architecture of the app. In the left part of the figure we have represented the student that uses the app, while in the left part we represented the author of the contents.

![Android app architecture](image)

Figure A.1. Android app architecture

The architecture is similar to the one implemented for the SG (presented in Figure 4.2), in which we decoupled the presentation, logic, and data layers. To test the EIVT, we embedded it in an Android application; some screen-shots of the app are shown in Figure A.2. The idea was that during the experiments, students could simply download the app from Google store and play with it.
We decided to start the application with a quest following an inquiry based approach. An inquiry based approach starts by posing questions, problems or scenarios rather than simply presenting established facts or portraying a smooth path to knowledge. For example, the initial quest proposed in physics insight was "Why the upper layer of the terrestrial atmosphere are colder?". When player logs in and starts playing, EIVT says "Hello, welcome to our app. How are you today?". The system analyses player sentence and replies something like 'I am happy to hear that you are fine' in case of positive answer or 'I am sorry that you are not fine' in case of negative answer. That is to sensitize the students, since the first usage of the app, to an emotional safe environment. Then, during the conversation, sentiment analysis is always performed on student sentences in order to catch any negative emotion and let the EIVT intervene with empathy. It is important to underline that the emotionally interlocution of the EIVT were oral. In fact, emotional behaviour was represented by the character’s animation (expression and verbal phrase) while the learning explanation were provided via text. That was our choice to reinforce the emotional aspect of the EIVT. In the app, the initial quest posed to the student is a sort of incipit that serves to encourage reasoning about a specific subject of study. Students ask free questions to the virtual tutor, designing their specific learning path and not following a predetermined one. That approach, even if commonly used by teachers especially in Anglo-Saxon countries, is rather innovative in gaming and educational game. That because, we argue, from an implementation point of view, it is much easier to have some, for example, quiz based approach rather than an inquiry based. An inquiry based approach based on free questioning implies a different organization of the supplied knowledge. In fact, the provision of content is not trivial for teachers (or content expert) because the content structure should be organised as question/answer pairs and it is rather important to foresee the largest number of possible pairs. That approach, tackles one very important aspect of ITSs, that is the shift from the programming of decisions to the programming of knowledge [106]. As already mentioned, the knowledge is structured as question/answer pairs. Teachers can create answers, and associate a set of possible expected questions to each answer. The authoring tools of others dialogue management systems are based on more complex formats, in order to specify rules where patterns of words correspond to a template that is executed when the player’s input matches the
corresponding pattern. Templates may implement recursive substitution (e.g. the AIML use of wild cards) which is very problematic. In particular if we consider the maintenance of the rules. These systems present databases with thousands of different rules, which make the writing of new ones or their adaptation to different contexts awkward for a non-expert user. The point is that rule-based systems are accurate and valuable for building realistic general-purpose dialogue system, but in the case of developing a system able to provide information to a student from a specific knowledge basis provided by a field expert, it seemed to us more appropriate to have a simpler method to provide the contents also for non-expert users (like teachers). We provided an authoring tool (Figure A.3), which is a web application where an author can register and, once verified, add his content. If a teacher desires to use the app to give a, for example, homework to his/her students, he/she can upload the desired content and ask the students to play with it.

Figure A.3. Authoring tool web interface
Appendix B

Learning contents

Every simulation room offers an interactive experiment’s panel that differs for each role, thus we have a total of 12 experiments. Each panel features a window in which the Virtual Tutor that explains to the player the assigned task appears. The list of tasks is reported in Table B.1.

<table>
<thead>
<tr>
<th>Role/Planet</th>
<th>Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physicist</strong></td>
<td></td>
</tr>
<tr>
<td>Maya</td>
<td>The player can modify the length of the string and the mass of the sphere that compose a pendulum, and see how they influence the resultant oscillation period. Inverting the pendulum formula he or she has to obtain the gravity on planet Maya.</td>
</tr>
<tr>
<td>Melissa</td>
<td>The player can modify the length and the angle of an inclined plane on which a sphere is let rolling, and see how the sphere’s final velocity varies. Inverting the inclined plane formula for the final velocity he or she has to obtain the gravity on planet Melissa.</td>
</tr>
<tr>
<td>Violet</td>
<td>The player can modify the mass of a rock hung up on a dynamometer, and see the correspondent weight force. Inverting the weight force formula he or she has to obtain the gravity on planet Violet.</td>
</tr>
<tr>
<td><strong>Astrobiologist</strong></td>
<td></td>
</tr>
<tr>
<td>Maya</td>
<td>The player has to observe a planet rotating around its star. When it crosses its equinox or solstice points, values from a table representing the average temperatures on the planet in that moment light up. He or she has to calculate the average temperature on planet Maya.</td>
</tr>
<tr>
<td>Melissa</td>
<td>The player recreates the Torricelli experiment, he or she can measure the height of the water or mercury-filled column. The atmospheric pressure on planet Melissa has to be calculated.</td>
</tr>
<tr>
<td>Learning contents</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Violet</strong></td>
<td>The player has to observe a planet rotating around its star. When it crosses its equinox or solstice points, values from a table representing the average day/night duration on the planet in that moment light up. The player has to calculate the average day/night duration on planet Violet.</td>
</tr>
<tr>
<td><strong>Chemist</strong></td>
<td>The player can look at the composition of the atmosphere of planet Maya with respect to the one of the Earth. Looking at the chart, he or she has to calculate the abundance of oxygen on planet Maya.</td>
</tr>
<tr>
<td><strong>Melissa</strong></td>
<td>The player can put a rock inside a beaker. When he or she does, the height of the water rises. Using the provided data, the volume of planet Melissa has to be calculated.</td>
</tr>
<tr>
<td><strong>Violet</strong></td>
<td>The player has to observe how water on planet Violet changes its state from liquid to solid to gas as a function of the temperature. He or she has to calculate at which temperature water evaporates on planet Violet.</td>
</tr>
<tr>
<td><strong>Geologist</strong></td>
<td>The player has to observe how the axis of planet Maya is inclined with respect to the Earth. Using trigonometric formulas with the provided data he or she has to calculate the axial tilt of the planet.</td>
</tr>
<tr>
<td><strong>Melissa</strong></td>
<td>The player can observe how the intensity of the magnetic field of planet Melissa varies as a function of the latitude. He or she has to calculate the average intensity of the magnetic field on planet Melissa.</td>
</tr>
<tr>
<td><strong>Violet</strong></td>
<td>The player can put a rock inside a beaker. When he or she does, the height of the water rises. Using the provided data the player has to calculate the mass of planet Violet.</td>
</tr>
</tbody>
</table>

**Table B.1.** Interactive experiments in the game
Appendix C

Additional work

The following publications are the result of an additional activity performed in collaboration with professors Luca Iocchi. We have investigated the effectiveness of robotics in education. Rather than creating excitement for children when playing with robots in games, we have examined the overall learning environment where a robot becomes a teaching assistant. We have designed two suitable lessons plan: one about the “Tower of Hanoi” problem and another one about "elastic hit" using a KUKA youBot as teaching assistant. The experiment involved two groups of students: one group was taught with the robot and the other group without the robot. We then compared the results of submitted questionnaire and the experience done with the different groups of students: we discovered that the youBot fits in with the existing robotics teaching challenging; it is highly motivating for participants and advances both academic and personal development skills. My role in this project has been to design the physics lesson plan, to give the lesson to the students, to prepare and analyse the questionnaires. Plus I have coordinated the work done with the schools (teachers and students).

The youBot experience has been preparatory to validate the assessment process used in real use-cases scenario.


Bibliography


Bibliography


Bibliography


Bibliography


