PlayPhysics:
Emotional Games Learning Environment for Teaching Physics

Karla Cristina Muñoz Esquivel, B.Sc. (Hons.), M.Sc.

Supervisors: Prof. Paul Mc Kevitt & Dr. Tom Lunney
External Collaborating Professor: Dr. Julieta Noguez Monroy (ITESM-CCM)

Confirmation Report

Date: 24th June, 2009

Presented as a requirement for Ph.D. in

School of Computing and Intelligent Systems
Faculty of Computing & Engineering
University of Ulster, Magee
Email: munoz_esquivel-k@email.ulster.ac.uk

Acknowledgements

I would like to begin by expressing my gratitude to my supervisors, Prof. Paul Mc Kevitt and Dr. Tom Lunney at the University of Ulster and the external collaborating professor, Dr. Julieta Noguez at Tecnológico de Monterrey, Mexico City Campus for their patience, experienced guidance and suggestions, since each of them has contributed much to my research and my life. Additionally, I would like to convey my thanks to the University of Ulster for supporting me with a Vice Chancellor's Research Studentship (VCRS). I want also to thank my colleagues at the Intelligent Systems Research Centre for their friendship, appreciation and support. Additionally, I would like to acknowledge, Dr. Philip Morrow, Head of Research Graduate School, and Dr. Joan Condell for the comments provided on my 100 Day Report and viva, which assisted me in refining this research proposal. I also would like to thank Patrick Kinsella, Ted Leath and Paddy McDonough for their technical support. Finally, I would like to convey my gratitude to the most important people in my life, my family, including my beloved boyfriend Dennis, thank you for always bringing me your support, believing in me and showing me that I can attain anything with perseverance and passion.

Abstract

Game based learning environments have proven to be inherently motivational. However, not all game based learning environments achieve required learning goals or motivate learning. These features can be achieved by following design principles and instructional design models, which support the curriculum, and by incorporating a new generation of Intelligent Tutoring Systems (ITSs). ITSs can intelligently guide, follow, assist and adapt to the learner's exploration, enhancing the learner's experience. New generation ITSs have the same challenges as older generation ITSs such as attaining the suitable domain and student knowledge representations and the adaptation, selection and communication of pedagogical responses. However, the new generation incorporates challenges of attaining an affective student representation and integrating affect in their pedagogical responses to enhance the student's learning and understanding. Attaining equilibrium between motivational, affinity and pedagogical strategies and communicating them unambiguously are still considered challenges. The new generation of ITSs arose when the area of Affective Computing, which aims to create machines that recognise and express affect, merged with ITSs, since affective processes proved to be deeply intertwined with cognitive processes, offering new possibilities to enhance the teaching-learning process. As a result, Affective computing has had an impact in student modelling, recognising the student's affect, modelling affective states and social interactions in synthetic characters and pedagogical agents. This research project aims to overcome the stated challenges by designing and implementing PlayPhysics, an emotional games learning environment for teaching Physics. The literature review discusses the state of the art of Affective Computing, Game based learning environments, ITSs and Knowledge representation.

From the comparison of PlayPhysics with related work, it is noted that PlayPhysics contributes to the state of the art in the research areas of ITSs, Game based learning environments and Affective Computing by providing a representation of interaction events and a student model to predict the learner's mental state (e.g., cognitive, affective and personal features). PlayPhysics also provides a coordinated, integrated and affective multimodal output involving the modulation of game features and characters. Olympia, the generic architecture of PlayPhysics, provides a reference for coherent patterns and abstractions and provides a framework to guide the system construction. PlayPhysics uses planning mechanisms that select affinity, motivational and pedagogical strategies to provide suitable feedback. ConceptNet, a common sense knowledge base, can be used to infer an affective valence from interaction events. Structured frames will be used to represent the interaction events. The Ortony, Clore and Collins (OCC) model will be used to identify elicited types of emotion. PlayPhysics identifies personality with goals, students' decisions and patterns of interactions based on personality theories. Dynamic Belief Networks will be applied to analysing the student's behaviour over time and influence diagrams, Markov decision processes or decision trees will be applied in providing intelligent planning mechanisms and selecting the pedagogical, affinity and motivational strategy or action. Software tools such as Haptek SDK, Torque Game Engine, Torque Game Basic, Unity Game Engine, Maya, Hugin Lite and ConceptNet, which can accelerate and support the development process of PlayPhysics, are being evaluated for suitability to the goals of the project. PlayPhysics will be evaluated with quantitative and qualitative questionnaires and the performance of the affective student model with personal traits against a model that only supports the student's affective features will be assessed using statistical analysis, such as linear regression. Finally, PlayPhysics aims to enhance the student's motivation, learning and understanding through the unambiguous communication of pedagogical, affinity and motivational strategies using the modulation of game elements, features and characters. A detailed explanation of how PlayPhysics will attain the stated features is provided in this report.

Keywords: Affective student modelling, Affective multimodal output, Dynamic Belief Networks, Educational games, Intelligent Tutoring Systems (ITSs), Intelligent planning mechanisms, PlayPhysics, Virtual Learning Environments (VLEs).

Contents

List of Tables

List of Figures

List of Acronyms

ACRONYM	MEANING
ABL	A Behaviour Language
ABT	Active Behaviour Tree
ADELE	Agent for DistancE LEarning
AI	Artificial Intelligence
ANNs	Artificial Neural Networks
API	Application Program Interface
ARCS	Attention, Relevance, Confidence and Satisfaction
ASIC	Application Specific Integrated Circuit
BPMS	Body Posture Measurement System
CAI	Computer-Assisted Instruction
DAG	Directed Acyclic Graph
DARPA	Defence Advanced Research Projects Agency
DBN	Dynamic Bayesian Networks
DGD1	Demographic Game Design 1
DIS	Distributed Interaction Simulation
DOF	Degrees of Freedom
ECA	Embodied Conversational Agent
EGEMS	Electronic Games For Education in Mathematics and Science
EMASPEL	Emotional Multi-Agents System for Peer to peer E-learning
EPA	Embodied Pedagogical Agent
ERA	Emotion Recognition Agent
ERPA	Emotional Response Predictor
ESTEL	Emotional State Towards Efficient Learning System
FIDGE	Fuzzified Instructional Design Development of Game-like Environments
FPGAs	Field Programmable Get Arrays
GUI	Graphic User Interface
IDDM	Instructional Design Development Model
IA	Intelligent Agents
HCI	Human Computer Interaction
IT	Information Technology
ITSs	Intelligent Tutoring Systems
LSA	Latent Semantic Analysis
MAS	Multi-Agents System
M3L	MultiModal Markup Language
MIT	Massachusetts Institute of Technology
MOCAS	Motivational and Culturally Aware System
OCC	Ortony, Clore, & Collins
OMCS	Open Mind Common Sense
P2P	Peer-to-Peer
PECS	Physical conditions, Emotional state, Cognitive capabilities, and Social status
RBF	Radial Basis Function
SDK	Software Development Kit
SIMNET	Simulator Networking
SVM	Supported Vector Machine
TTS	Text-to-Speech
UML	Unified Modelling Language
VLEs	Virtual Learning Environments

1 Introduction

Students' attitudes, aptitudes and expectations have evolved to adapt to the Information Technology (IT) revolution and applications with high media content (Oblinger, 2004). The technological acceleration and advances in artificial intelligence (AI), neuroscience, psychology and cognitive science have resulted in dramatic changes and promises in the computer tutoring field. To assist learning in an intelligent manner, research has focused on addressing computer perception, the representation of the domain, the representation of the student's knowledge and emotions and the adaptability and sequence of responses, e.g. activities, problems and feedback. The key question has been to what degree do systems have to imitate human teachers (Du Boulay & Luckin, 2001; Picard et al., 2004).

Currently there are several research groups focused on the creation of a new generation of intelligent tutoring systems (ITSs) (Conati & Maclaren, 2009; D'Mello et al., 2008a; Sarrafzadeh et al., 2008; Chaffar & Frasson, 2004; Chalfoun & Frasson, 2008). These ITSs promise an enhanced awareness of the student's knowledge, understanding, personality traits and emotional disposition. Accordingly, the ITSs can support and improve the learning process by encouraging the student's motivation for learning, showing affinity, providing accurate guidance or optimising the learner's emotional state for improved understanding. The key challenges involved in this research are how an ITS can acquire and represent the intellectual and emotional state of the learner and also how an ITS can communicate with the student for enhancing motivation and facilitating deep understanding.

Recognising a person's affective state is a challenge even for another person. Methodologists in social and behavioural sciences noted that a human expert recognises emotion with 75% accuracy (Robson, 1993). This limitation is due to diverse factors, such as differences in personality traits, cultures and social standards and rules. Also, emotions occur in mixed forms. Computers are attempting to recognise and produce affect (Picard, 1995). In attaining this objective physiological signals, body language, facial gestures, prosodic and acoustic features of natural language, written natural language and selection of colours are employed. The computer systems that employ hardware devices to acquire the user's affective features only recognise the visible results of affect and not their causes (e.g. D'Mello et al., 2008). Additionally, this approach cannot be employed in online applications, since the required processes of classification reduce the performance of systems and if hardware is to be used on the client side, it is not guaranteed that the student possesses it. Other approaches attempt to predict the user's affective state from the sources that can elicit it. These approaches use goals, personal traits and social standards to identify the activation of the affective state. Chalfoun et al. (2006) have claimed the success of this technique, even though the work has not been validated. This approach can be used in e-learning to enhance the student's knowledge, understanding and motivation. The prediction of the affective state of the learner can be deployed in implementing new generation ITSs. In addition to the challenge of implementing a knowledge representation of the learner's cognitive state in an ITS, this new generation has the challenge of attaining a flexible, accurate and effective knowledge representation of the learner's personality traits and affective behaviour (e.g. Conati & Maclaren 2009). It is important to note that the learner's knowledge representation influences the pedagogical and motivational strategies that can be implemented. Additionally, research in this field is focused on solving the problem of determining which interaction information must be considered to recognise or predict the affective state of the learner (e.g. Conati & Maclaren 2009 and D'Mello et al. 2008).

Game based learning environments have proven to be highly motivational learning tools (Sykes, 2006), but this is dependent on their architecture and design, since not all game learning environments result in student learning (Blanchard & Frasson, 2006). Game based learning environments can obtain multiple advantages by including ITSs in their architectures, since ITSs can help guide, follow and assist the student during the learning process, enabling the student to attain the learning goals (Noguez & Huesca, 2008). The new generation of ITSs have the objective of using affect to enhance the student's motivation for learning. As a result, these ITSs have to implement pedagogical and motivational strategies. Attaining the balance between both strategies is still a challenge (Du Boulay & Luckin, 2001; Sarrafzadeh et al. 2008).

In addition, educational games manage multiple media, e.g. sounds and graphics, to attain a high level of interactivity, hence providing immediate feedback to the learner's actions (Sykes, 2006). These media have diverse affective features such as colours, audio and acoustic tones and game characters that can be leveraged to communicate learning and motivational messages. The game story is also a way of handling the student's attention and level of engagement as was depicted by Dias & Paiva (2005). In a process of communication when information is not employed intelligently and appropriately to support the message, the message is not understood and the information is perceived as noise. This problem is also present in the state of the art of pedagogical agents and synthetic characters, where non-verbal language is used in combination with verbal language to communicate a message. Therefore, it is noted that enhanced techniques of content synchronisation, integration and planning are necessary to attain an intelligent management of media sources and communicate the intention of the system unambiguously.

ConceptNet is a common sense knowledge base useful for inferring affect from independent clauses, even when the clauses do not contain affective terms (Liu & Singh, 2004). Hence, common sense can facilitate inferring the affective valence of interaction events. The Ortony, Clore and Collins (OCC) model is a cognitive theory of emotion, which classifies emotion according to the way that it was elicited and the person's standards, attitudes and goals (Ortony et al., 1999). Also, Dynamic Belief Networks, influence diagrams, Markov decision processes and decision trees are potential Artificial Intelligence (AI) techniques, which may enable the attainment of suitable student representation and planning mechanisms (Woolf, 2009). Planning mechanisms enable the selection of the most suitable pedagogical, motivational and affinity strategies, which optimise student knowledge and understanding. Based on these ideas and the stated challenges, this research proposes PlayPhysics, an emotional games learning environment that incorporates a new generation ITS. The key aim of PlayPhysics is to enhance the student's level of engagement, learning and understanding through the intelligent selection of motivational, affinity and pedagogical strategies. These strategies will be unambiguously communicated by modulating the game elements, features and characters. Accordingly, the potential contributions of this research project proposal are, a representation of interaction events and a student model to predict the learner's mental state, an integrated and affective multimodal output involving the modulation of game features and game characters, and planning mechanisms that select affinity, motivational and pedagogical strategies. PlayPhysics will be developed with the Olympia architecture, which is a framework to guide the system construction. PlayPhysics will be designed, implemented and evaluated with specific case studies of teaching Physics at undergraduate level.

Section 2 focuses on a literature review of research in tutoring systems such as Virtual Learning Environments (VLEs) and Educational Games, Intelligent Tutoring Systems (ITSs), Affective and Educational Applications and Knowledge Representations. A summary of the challenges, found in these areas, is provided. Section 3 discusses the research project proposal with its aims and objectives, rationale, hypotheses and methodology, the analysis of requirements, the Olympia architecture to be used in the PlayPhysics application. A comparison of PlayPhysics with previous work is provided by highlighting its potential research contributions. Section 4 discusses the potential software tools for developing PlayPhysics. Section 5 discusses the research project schedule. Finally section 6 concludes and discusses future work.

2 Literature review

The introduction of flight and space simulators in the USA indicated the origin of VLEs and educational games (Kincaid, 2006). In 1952, the first computer games appeared. Some of them were computer-based military simulation games, which were developed by The Rand Air Defence Lab in Santa Monica. For security reasons, these games could only be operated by researchers and scientists. In this period, the cost of computer time was notably high and the number of computers was considerably limited (Bergeron, 2005). In the 1980s and 1990s, the Defence Advanced Research Projects Agency (DARPA) started the Simulator Networking (SIMNET) program. Its results were the design of important simulation interaction principles and the creation of a messaging protocol to exchange data. This protocol was the predecessor of the Distributed Interaction Simulation (DIS) protocols (Kincaid, 2006). In 1997, Commercial games and their joint research with the entertainment industry were ahead of the USA's Department of Defence simulations. In 2000, the information technology (IT) revolution and a new reform spirit in education began (Steels & Tokoro, 2003). Simulators started to merge with commercial video games (Kincaid, 2006). The Woodrow Wilson International Center for Scholars founded the Serious Video Games Initiative in 2002. The aim was to announce a new series of policies focused on the development of education, exploration and management tools utilizing state of the art of computer game design (SGI, 2008). Since then, virtual learning environments (VLEs) and educational games have been in a continuous process of evolution. The aim is to enhance learning by meeting the students' expectations under specific constraints.

Virtual laboratories are related to a specific thematic discipline and represent an educational approach in which the student is able to interact with equipment or perform experiments in a remote or simulated way. In the latter case, the student interacts with computational models of the equipment and its environment (Noguez & Huesca, 2008). In a remote approach, the laboratory equipment is visualised as software entities that are comprised of a well-defined interface with a specific functionality. The objective is to enable the configuration and management of the devices via online (Czekierda & Zielinski, 2006). An educational game is an interactive computing application. It has challenging entertaining and learning goals. It uses concepts of scoring, winning or losing. A skill, knowledge or attitude applied in real life can be taught through educational games. An educational game may or may not have significant hardware components (Bergeron, 2005). VLEs originated through the combination of intelligent tutoring systems (ITSs) with virtual laboratories. ITSs are coupled with virtual laboratories or educational games to enhance the students' learning and understanding. Traditionally, ITSs have four main modules: the domain model, the student model, the teaching model and the graphical user interface (GUI). The ITS' modules can contain AI techniques, which use representations of the domain knowledge to understand the student's behaviour and provide an intelligent response (Freedman, 2000). ITSs can provide advantages in the student's exploration and interaction through following the performance of the student in real time, using task planning to enhance the student's learning and providing feedback, help and guidance to address the student's immediate problems (Noguez & Huesca, 2008). A new generation of ITSs has arisen to meet the requirements of an enhanced awareness of the student and the provision of effective pedagogical responses. The aim has been to be as effective as human tutoring (Sarrafzadeh et al., 2008). Research in the area of affective computing has contributed considerably to the attainment of this goal (Picard, 1995). Affective computing, which aims to enable machines to recognise and express affect, has had an impact in student modelling (Conati& Maclaren, 2009), recognising the student's affect (D'Mello et al., 2008a; Sarrafzadeh et al., 2008), modelling affective states and social interactions in synthetic characters (Dias et al., 2006) and pedagogical agents (Marin et al., 2006). This section is focused on discussing the challenges, main advantages and disadvantages of VLEs and educational games. Design principles and an instructional design development model (IDDM), which can assist in the creation of game based learning environments are introduced. Also, the main developments and challenges associated with ITSs are overviewed. Also, other affective and educational applications that can progress the research area of ITSs are discussed. Finally, some knowledge representations are described and a summary of the research challenges, related to these areas, is provided.

2.1 Comparison of traditional education with computer tutoring

VLEs and educational games have advantages and disadvantages when compared with human tutoring. The knowledge delivered in traditional tutoring depends on the subject taught, the experience of the lecturer and the lecturer's knowledge of the world. Alternatively, in VLEs and educational games the knowledge is what the student needs to learn. In addition, students acquire skills to manage the available tools. In traditional learning, a specific time and place is agreed for the learning. Using VLEs the student has more flexibility, since the student can decide where and when to learn. In an educational game the place depends on the targeted game platform (e.g. online or console games), but the student can decide when to learn. Traditional education has difficulty adapting to each student's rhythm and way of learning, while educational games and VLEs adapt easily. VLEs and educational games enable the student to immediately apply the acquired knowledge. In traditional education, the student usually has a spectator role. The student's role is only interactive when the student's learning subject is related to experimentation. The human tutor can give suitable feedback in a classroom, while in a VLE and educational game, the response depends on the available media, resources and implemented intelligence mechanisms (Slator et al., 2002; Bergeron, 2005; Hsiao, 2007). Summarising, VLEs and educational games are interactive, provide personalised attention and evaluation, encourage the acquisition of experience and abilities and reduce the teaching cost per student. Sykes (2006) outlines the games' abilities of influencing the student's emotional state and offering immediate feedback to the student's actions. Games incorporate storytelling, which has played an important historical role in the way that humans learn.

Additionally, in one-to-one human tutoring, the tutor adapts to the affective state of the learner in different ways. It is not just about presenting the most suitable pedagogical action. It is about being believable and spontaneous. In one-to-one human tutoring, answering questions is the most frequent student action. In the study executed by Alexander & Sarrafzadeh (2008), results showed that question-answering actions occurred 92% of the time. Also, tutor actions occurred approximately 71% of the time, such as asking questions, querying for additional information and delivering positive and neutral immediate feedback. Human tutors almost never used negative feedback. The study was carried out to determine the behaviours and facial expressions during the interaction between tutors and students while teaching Maths. The study, reported that students and tutors show neutral face expressions 86% of the time. During the teaching-learning process, the student's expressions influenced the tutor's expressions 90% of the time. Smiling expressions were observed 29% of the time in students and 12% of the time in the tutors with high and low intensity respectively. 4% of the time students appear apprehensive and 3% of the time confused. Other facial expressions in tutors and students occurred 2% of the time. In this study, it was noted that the tutor's responses are clearly influenced by the history of interaction. As a limitation of this study, it is important to note that the affective acoustic and prosodic information corresponding to the tutor and lecturer was not captured. Also, in a similar study executed by Ghijsen et al. (2005), which attempted to recognise the students' mental state from their facial expressions, the predominant facial expressions between students of nursing and tutors were neutral and smiling.

Pearson & Graesser (2008) carried out a study observing pedagogical strategies, dialogue patterns and tutor feedback. They noted 14 facts about human tutoring. Ideal tutoring models are rarely adopted. The most common responses are persistent questions, hints, elaborations and linking ideas. Immediate feedback is preferred to delayed feedback. Varied feedback is provided to student's errors. Error-specific feedback to provide remediation is low. Student errors are never attributed to lack of ability. A question is considered the primary unit of communication between student and tutor. Questions confirming the validity of the student knowledge are predominant. Deep-reasoning questions are predominant in high-quality students. Answers are used to assess the student's understanding. Examples are commonly used during human tutoring to explain difficult material. The material is mostly concrete and difficult and frequently comes from curriculum scripts and textbooks.

Challenges of VLEs	Common challenges	Challenges of educational games
• From an industrial perspective: The trade-off between self-development and work productivity. • Decide the learning approach that will be implemented (e.g. Problem Based Learning). • Attain adaptability to each student. • Choose the intelligent tutoring technique and the suitable pedagogical action. • Acquire, exploit and maintain knowledge models	• Assess the necessary requirements of infrastructure, such as bandwidth and telecommunications. • Attain the required software quality. • Address the life-long learning requirements.	• The natural caution to adopt gaming as a learning aid. • The existing scepticism to attain the learning goals. • The mismatch between the game content and the educational curriculum.

Table 1. Challenges of VLEs and Educational Games

2.1.1 Challenges of educational games and VLEs

Educational games and VLEs have to overcome several challenges in order to be adopted as reliable sources of learning. These challenges are summarised in Table 1 through analysing the work of Du Boulay & Luckin (2001), McFarlane et al. (2001), Slator et al. (2002), Liu et al. (2002) and Carpenter & Windsor (2006). In addition, Educational games easily attain the student's interest, but it is difficult to ensure that the student is thinking in the domain knowledge when learning to play a game successfully (Conati & Maclaren, 2009). VLEs have proven teaching and learning effectiveness, but engaging the student in the learning process is still a challenge (Kuo, 2007). For that reason a mixed approach, which synthesises features of both learning environments, may lead to enhanced results (Conati & Maclaren, 2009; Muñoz et al., 2009b).

2.1.2 Educational games design

To limit the influence of the challenges stated in section 2.1.1 educational games have to follow design principles and instructional design development models, which ensure their reliability when attaining learning goals. The Games-to-teach project was created by the Massachusetts Institute of Technology (MIT) and Microsoft with the objective of creating games that support learning. The Games-to-teach research team signalled design principles to create educational games (Squire et al., 2003). The team suggested beginning the creation of an educational game from a standard simulation, using objects, i.e. intrinsic motivators, to enhance the abilities and basic attributes of game characters, setting the context of the educational game according to an identified contested space, exploring the use of temporal transgression experiences, e.g. breaking social boundaries and performing new roles, as engaging features and identifying real-world applications of the concept that is going to be taught. In addition, the Games-to-teach research team recommended using goal based scenarios (Schank, 1994) and designing decisions, consequences and joint goals (Squire et al., 2003). The Fuzzified Instructional Design Development of Game-like Environments (FIDGE) model is an IDDM aware of the real-world imprecision, since it was originated from data collected from real-life practices. The FIDGE model consists of pre-analysis, analysis, design, development and evaluation phases. Each phase has fuzzy boundaries and the progression between them is non-linear. This model is mainly focused on managing time, taking early decisions about the utilities and limitations of the used technology, obtaining support from literature reviews, performing continuous evaluations and attaining flexibility and modularity in the final product (Akilli & Cagiltay, 2006). The First Principles design method enables a long view of the design process, since firstly the goals are set and the nature of the game world abstraction is determined. After this, it is possible to proceed to the design and implementation of the game (Bateman & Boon, 2006). In addition, at the design stage of a serious game, it is very important to adapt the gameplay to the audience and sort the audience correctly, since it improves the reception of the game. To achieve this objective, the Demographic Game Design 1 (DGD1) model can be used. DGD1 is a tool to develop games, which incorporates elements designed to meet the players' needs and was defined using as reference Myers & Briggs personality theory. DGD1 identifies four play styles or preferences while interacting and learning defined as: *Conqueror, manager, wanderer and participant* (Bateman & Boon, 2006).

2.2 Intelligent Tutoring Systems (ITSs)

ITSs can assist students interacting with VLEs and educational games to attain specific learning goals. ITSs, highly related to Cognitive Science, were originated in 1970s. The difference between an ITS and Computer-Assisted Instruction (CAI) was that an ITS used new AI formalisms to isolate the domain knowledge data from the logic of interaction with the student. Advantages of the ITSs approach are to follow the student's performance, generalise and apply the logic of teaching to different problems and domains and to use a student model as a reference for tutoring. The latter is inferred from the student's behaviour (Clancey & Buchanan, 1982). In order to advance in the ITSs area, research has concentrated on the representation of the domain knowledge, the representation of the student, the adaptation and selection of pedagogical responses and their communication (Du Boulay & Luckin, 2001). To attain the representation of the student, abstract models of what the student has to learn, the learning process and how emotion is aroused in the student have been implemented. The definition of the student model is related to the pedagogical strategies that are going to be implemented, since it is assumed that the student's errors reflect misconceptions about procedures or facts in the domain knowledge (Clancey & Buchanan, 1982). Recently, the dimension corresponding to the student's personal traits has been integrated into student models such as in Conati & Maclaren (2009). The objective is to attain understanding of the student's behaviour and mental state.

Teaching involves communication actions such as explaining, answering, persuading, asking, demonstrating and describing. Initially, the pedagogical techniques implemented in computer tutoring were isolated skills (Du Boulay & Luckin, 2001), since these ITSs did not have the necessary abilities to communicate their responses with the features depicted by human dialogues such as the LISP tutor of Anderson & Skwarecki (1986). Anderson & Skwarecki (1986) attempted to implement all the possible responses of the tutor to each answer of the students, which resulted in a production system of hundreds of ideal and buggy rules. The latter were attached to natural language templates to provide feedback. LISP tutor used a Model-Tracing methodology, where the tutor monitors the student constantly to provide guidance while the student is solving specific problems. LISP tutor provided immediate feedback, but there are cases where the delay of feedback is necessary to enable students to become independent in their learning. As a result, it was noted that the use of more intelligent teaching strategies and more improved dialogue capabilities had to be attained to enhance learning.

GUIDON (Clancey & Buchanan, 1982) and MENO tutor (Woolf & McDonald, 1984) were pioneer ITSs that incorporated communication skills. GUIDON focused on finding ways of representing teaching and problem-solving knowledge, managing them separately. The targeted problems were how to recognise the student's knowledge and how to explain the domain knowledge. GUIDON establishes a dialogue with the student to discuss a medical or engineering case. The MYCIN knowledge base was used to encode production rules. The rules were used to provide pedagogical responses. Students learn through testing their hypothesis. The interaction of GUIDON was based on Socratic Tutoring (Collins et al., 1975), which is comprised of a series of questions that assist the student in identifying misconceptions and knowledge that needs to be acquired. MENO tutor used natural language dialogue to provide the students with feedback using different tutorial methods and strategic rules such as asking questions and giving explanations about specific topics. Woolf & McDonald (1984) focused on solving separately the problems of generating a pedagogical plan and a tutoring dialogue. A communication module in an ITS architecture was suggested, which defines syntaxes and rhetoric features of the dialogue (Evens & Michael, 2005; Du Boulay & Luckin, 2001; Woolf & McDonald, 1984). Communicating the pedagogical responses of ITSs involves the challenges of determining how messages are received and understood and how answers have to be formulated. To attain more effective communication it is necessary to look at the semantic meaning beyond the words, since the student's input is still not completely understood. MENO tutor reviewed earlier dialogues with students and adapted its dialogue appropriately. In the 1980s, ITSs had an irregular progress in areas such as provision and implementation of pedagogical actions and strategies, development of communication skills and implementation of theories of motivation and affect. In that period trying to incorporate into computer tutoring the skills and abilities of human tutors required AI techniques that were beyond the state of the art (Du Boulay & Luckin, 2001).

Ohlsson (1987) set the scene of the challenges that ITSs have to overcome in order to attain the effectiveness of human tutoring. The student's goals, cognitive competence, learning style and learning rate are student characteristics that may be considered by ITSs to attain adaptable instruction. Strategies that link teaching goals to actions are needed in ITSs to generate adaptable education. ITSs have to adapt the form of instruction and the content to attain the student's understanding. The domain knowledge can be represented in different ways. Therefore, the learning acquisition has many valid end states. ITSs must attain a diagnostic of the student's knowledge and understanding as part of their teaching plan. This diagnostic is attained through comparing the student's performance with the ITS expectations represented by their student model. The teaching plan of an ITS must consider the representation of the student, the student's performance, the learning goals and the relation between student's actions and learning goals. An adaptable ITS will be able to choose an action from a wide set of pedagogical actions based on more information than the last student's interaction. To increase the performance of ITSs, teaching strategies must identify the circumstances when a specific learning tactic will attain the maximum gain of student's understanding and motivation. In addition, to attain the student's reflection of the domain knowledge complexity, voice and gestures or other media, such as graphics and sounds, can assist in revealing its essential meaning.

In the 1980s there were restrictions of modality that limited the attainment of the teaching and learning goals, hence text and graphics were the most common resources (Du Boulay & Luckin, 2001). Also, the student's meta-language, mental representations of the domain knowledge that the student uses to evaluate, infer or create new knowledge, is more difficult to constrain than the student's natural language related to the domain knowledge. As a result, the ITSs of that time were less focused on attaining reflection and representing the student's meta-cognition. Nowadays, that perspective has changed and ITSs are focused on implementing diverse strategies that attain educational interaction and a more learner-centred approach, which encourage meta-cognition skills such as self-evaluation and self-explanation (Du Boulay & Luckin, 2001). In addition, new AI techniques and knowledge representations, such as semantic networks, probabilistic models over time and planning mechanisms, have arisen between the 1990s and the 2000s, which have contributed to overcoming some of the stated challenges (Russell & Norving, 2003; Sucar & Noguez, 2008).

2.2.1 Development approaches for Intelligent Tutoring Systems (ITSs)

There are three main approaches for developing ITSs: (1) The observation of human tutoring in its cognitive and affective aspects; (2) the analysis of learning theory; and (3) the observation of students when attempting to solve specific problems and inferring knowledge. The first approach can be used for the implementation of rules, the second is used to derive theories and the third is to attempt a solution through the implementation of student models. A detailed explanation and review of these approaches can be found in Du Boulay & Luckin (2001). Significant principles originating from this review are:

- A combination of ideal or standard pedagogical strategies may attain effective ITSs.
- Human-teachers use goals, beliefs and knowledge that are context-specific.
- The effectiveness of the pedagogical responses depends on the human teachers' personality traits.
- To attain teaching equilibrium it is important to provide feedback considering the tactical and strategic perspectives. The former is preferred when immediate understanding is pursued and the latter when the conformation of independent learners is explored. The excess of any of these techniques can result in the inhibition of problem-solving skills in the learner or the feeling of being lost and de-motivated.
- The GUI can aid the student in understanding the nature of the problem-domain.
- Affective issues aid students in maintaining their confidence, curiosity and sense of control over the learning situation. It is a way to be indirect when setting goals and reacting to errors.
- The student effort is also evidence of self-motivation. Therefore even when the performance is not optimal, learners deserve to be praised. The use of extrinsic rewards can also be managed to attain learning motivation.
- Some of the pedagogical strategies implemented in ITSs are Socratic Tutoring, Conversation Theory, Contingent Teaching, Fine-grained diagnosis and remediation, Procedural knowledge, Constructionist Theory, Reciprocal Teaching, Self-regulated learning and Model Tracing.
- The fidelity of the simulated environment plays a major role in making predictions about the student's behaviour.

Managing the motivation to learn is a goal of ITSs. Instructional theories of motivation were introduced by Malone & Lepper, (1987) and Keller (2006). Malone and Lepper (1987) set the view of education over educational games, which are intrinsically motivating. Challenge, fantasy, curiosity, control, cooperation, competition and recognition are proposed factors and design features that can attain the student's motivation. Keller (2006) introduced the Attention, Relevance, Confidence and Satisfaction (ACRS) motivational design model, which is grounded in expectancy-value, reinforcement and cognitive evaluation theories. ACRS model is a problem-solving approach to identifying motivational constraints in the teaching-learning process. Its strategies are related to the course and material design and teaching style. The ACRS model explains the relation between effort, performance and satisfaction. According to this theory, learners are motivated when they are *attentive*, show interest for the *content* and are *reasonably* challenged. However to keep the student's motivation, *satisfaction* must be accomplished from the learning process. A significant contribution of the work of Keller (2000) is the observed curvilinear relationship that exists between performance and motivation, which is shown in Figure 1. As can be seen, while motivation increases, performance increases until reaching an optimal point. Afterward, motivation increases while performance decreases. This phenomenon is due to stress, since a degree of stress or tension is always associated with motivation. Before attaining the optimal point the stress receives the name of *facilitative* stress and after the optimal point the stress receives the name of *debilitating* stress.

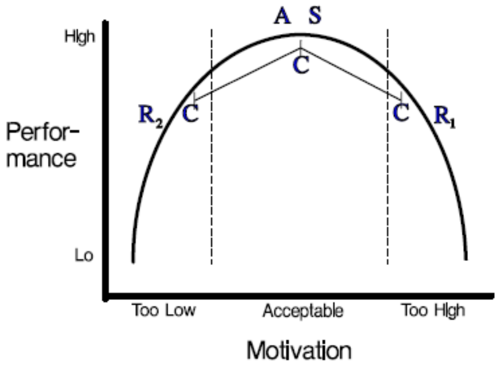

Figure 1. Relationship between performance & motivation (Keller, 2000, p. 7)

The motivational features of ITSs are attained mainly by using the two approaches of reasoning about the affective states of human tutors and also the affective states of students. The former involves the creation of pedagogical agents such as in Marin et al. (2006) to attain the student's engagement. In the latter approach, ITSs must include in their infrastructure a knowledge domain representation and a student model, which will support the recognition of the current student's motivational state (e.g. Conati & Maclaren 2009) and will react to attain an optimal state (e.g. Chaffar & Frasson 2004). The instructional and motivational plans in ITSs can be handled separately or together. In the case that the two plans are handled together, distinguishing the cases where motivational strategies compete with educational strategies is essential. The challenge of applying both pedagogical and motivational plans is to attain flexibility and adaptability between them and the student's goals and actions (Du Boulay & Luckin, 2001). These challenges may be addressed by following motivational design theories in combination with planning strategies. Hence, the addition of motivational features and plans to ITSs has set the perspective on attaining a new generation of ITSs, which in addition to the features and challenges of their ancestors have the challenges of recognising and stimulating affect to attain an enhancement in the student's understanding, learning and engagement.

2.2.2 A new generation of Intelligent Tutoring Systems (ITSs)

The addition of the emotional dimension to ITSs promises interesting advantages in the learning process such as an enhanced motivation, interest and attention (Picard, 1995). Human cognition and emotion are deeply interrelated. Cognitive mechanisms interpret, understand, analyse and remember elements of the world while affective and emotional mechanisms evaluate and judge. Hence, affect modulates the operative parameters of cognition, e.g. weights and thresholds, that bias the cognitive mechanisms, influence the level at which processing takes place and the locus of attention. Therefore, negative affect when accompanied by high arousal seems to lead to more focused and deep processing and positive affect seems to lead to broad and widely spread processing. In the former case, a human experiences increased vigilance (i.e. stress) and in the latter the human experiences increased curiosity (i.e. creativity) (Norman et al., 2003).

Research has been focused on the creation of a new generation of ITSs, which aims to recognise and provide affect in order to enhance the teaching-learning experience, hence improving the student's understanding and personal disposition to learn. Recognising affect has been the main challenge of this new generation of ITSs and until today no autonomous system can reliably recognise all the emotions during learning. To recognise the student's affective state, researchers have followed three approaches: Identifying the physical effects of the learner’s affective state (D'Mello et al., 2008a), predicting emotion from its origin (Chalfoun et al., 2006) and a combination of the previous two approaches (Conati & Maclaren, 2009). Additionally, these systems have the aim of using motivational strategies to enhance learning and change the learner’s affective state to the one optimal to attain knowledge and understanding. This section is focused on giving a brief discussion and overview of the ITSs that address the recognition of emotion using one of the three stated approaches, their aims, objectives, pedagogical strategies and challenges.

Autotutor

Autotutor is an ITS used in VLEs for teaching Newtonian Physics, computer literacy and critical thinking at undergraduate level. It recognises the moods of boredom, engagement, confusion and frustration through monitoring communication features in dialogues, body language and facial gestures. Autotutor aimed to overcome the recognition of affect with real-time constraints through sensors and signal processing algorithms and the selection of pedagogical actions that enhance the student’s motivation and learning. Autotutor interacts with the student through written natural language. Challenging problems were implemented. The Autotutor’s feedback is comprised of direct questions, evaluations of the student’s written explanations and answers, hints and answers to the student’s questions. Additionally, it summarises, identifies and corrects misconceptions and asks the student to fill in missing information with assertions (D'Mello et al., 2008a).

The techniques utilised to classify the student’s affective features were supervised learning methods. The training data corresponded to 28 participants and was collected through sensors during the students’ interaction with Autotutor. The affect labels were annotated manually through the judgments of students, peers and trained judges. To assess communication cues it is necessary to take into account assessment of deep meaning, world knowledge (i.e. common sense) and pragmatic aspects of communication. In addition, it requires essential knowledge about the domain. To overcome these challenges, conversational features and discourse markets were extracted from the Autotutor’s log files, which included temporary features, assessments of response verbosity, quality assessments of the student’s response using Latent Semantics Analysis (LSA), conversational tutoring guidance and feedback. LSA is a natural language processing technique used to analyse statistically the relationships between a set of documents and their contained terms. LSA produces a set of concepts related to both, documents and terms (Landauer et al., 1998).

The body posture measurement system (BPMS) was used to assess the student’s body posture and to monitor the facial features the IBM BlueEyes system was used. To recognise the facial emotion, muscles in the face were labelled and linked to an emotion (Ekman & Friesen, 1978). The BlueEyes system attains 65% accuracy. Searching for the most effective classification algorithm, seventeen standard classification algorithms that use k-fold cross validation were used to assess the conversational clues (D’Mello et al., 2008a). Some of the classification algorithms were Bayesian classifiers, functions, instance based techniques, meta-classification schemes, trees and rules (D’Mello et al., 2008b). These algorithms classified boredom, confusion, engagement and frustration with 69%, 68%, 71% and 78% accuracy. Also, gross body language was classified with 70%, 65%, 74% and 72% accuracy respectively. The neutral base line is 50%. From a global perspective, Autotutor classifies the affective states with 73% accuracy (D’Mello et al., 2008a).

Autotutor focused on changing negative affective states (e.g. boredom, frustration and confusion) to the ones that promote flow during interaction (i.e. engagement).The challenge is to enable the ITS to intervene without being perceived by the students as an obstruction to their interaction. To address the negative states, D'Mello et al. (2008a) used theoretical foundations of pedagogy and affect and the suggestions of pedagogical experts. The two theoretical foundations used were attribution theory (Heider, 1958) and cognitive disequilibrium (Piaget, 1952). Empathy was also used by Autotutor to modulate its response, since it is an anticipation mechanism, where the observer has a feeling of participation in the experience of the other through understanding his or her feelings and intentions (Bischof-Köhler, 2004). From the pedagogical strategies, production rules were derived to provide feedback. The feedback was provided through dialogues with the embodied pedagogical agent (EPA). The EPA’s voice is modulated through a text-to-speech engine using variations in pitch, speed rate and other prosodic features. The expressed emotions are surprise, delight, disappointment, compassion and scepticism (Ekman & Friesen, 1978). The student model of Autotutor corresponding to its cognitive state is represented through fuzzy production rules.

It was noted that the extracted features of the dialogue do not ensure by themselves an accurate identification of affective states (D'Mello et al., 2008b). Autotutor encourages the students' thinking skills, but the interaction with Autotutor is only through written text, the students are not able to learn by doing, only by examining the validity of their ideas and knowledge. In addition, the written text did not allow the identification of affective acoustic and prosodic features. The challenge involved in recognising emotion is to deal with data that is incomplete and not understood. The data is also influenced by individual differences that Autotutor has not attempted to understand. The Autotutor's assessments are not error free or complete, which shows that the LSA scores require improvement in semantics and cohesion between terms (D'Mello et al., 2008b). Additionally, it is important to note that the use of hardware to recognise emotion has proved invasive, hence making students uncomfortable. The use of hardware on the students' side is not effective for e-learning, since students have to be transported to special research labs where these ITSs are located. Occasionally faults in the technological equipment are an obstacle to the student's interaction (Burleson & Picard, 2007). Clore & Ortony (2000) stated that it is important to distinguish between moods and emotions, since from a psychological and learning perspective, emotion has implications for managing that moods do not. Moods are just feeling states, which can arise from completely physiological causes, i.e. the information is not necessarily feedback about the current situation, and can influence the intensity of an emotion. ITSs, which are comprised of advanced technological equipment, attempt to recognise the visible effects of emotions that may be deeply intermixed with moods, but from the nature of the latter, ITSs cannot ensure that the features that are identified are originated during the learning experience. In addition, D'Mello et al. (2008a) have still to work in the believability of the EPA, e.g. the avatar has a distinctly inhuman stare and the management of prosodic and acoustic features can be improved to depict affect more accurately.

PrimeClimb

PrimeClimb, created by the Electronic Games for Education in Mathematics and Science (EGEMS) group, is an educational game for teaching maths (e.g. number factorization) to students in an age range between 10 and 12. In PrimeClimb, the student's emotions are recognised using a student model implemented with Dynamic Bayesian Networks (DBN). This approach utilised the Ortony, Clore and Collins model (i.e. OCC model). This cognitive theory of emotion defines types of emotion according to their origin. Events, agents and objects can elicit a specific emotion. Emotions are influenced by personality traits and cultural context. As a result, events are evaluated against goals, actions against social standards and objects against attitudes. Each type of emotion is influenced by specific intensity variables (Ortony et al., 1990) and has a threshold value. When the intensity of the emotion is above the threshold value, the emotion was activated.

In PrimeClimb, a DBN involving the causes (i.e. state of interaction) and effects (i.e. body expressions) of emotion during learning, shown in Figure 2, was defined (Conati & Maclaren, 2009). PrimeClimb aimed to enable the EPA to intervene in the student's learning process and to enhance the student's understanding and motivation. The main challenges involved when using the stated DBN technique are the creation of mechanisms and techniques to assess the student's goals and the process of evaluation used to know the likelihood of attaining them. Also, deciding which causal information has to be saved and gathered from intrusive and non-intrusive interaction events is another problem to overcome. Additionally, it is necessary to know if the use of intrusive methods, which are generally more expensive, is necessary. Knowing why an emotion was elicited brings more information and more possibilities to change negative emotions. Additionally, empathy can be expressed by an EPA to show a more believable behaviour. This approach is focused on identifying immediate emotions, which evolve during time until they become long-term affective states or moods.

Conati & Maclaren (2009) used visible interface outcomes originated through the student and EPA's actions to analyse the student's level of arousal. The desirability of the event, an intensity variable of emotion, depends on the student's personal traits, which Conati & Maclaren (2009) associated with the personal goals of each student while playing the game. Assessing these goals is another difficult challenge, since asking directly what the student’s goal is while interacting can be perceived as intrusive. Hence, student’s interaction patterns and the outcome of actions were used to propagate evidence and identify the student’s goals. Goals were represented as Boolean variables that did not change over time to limit the complexity of the approach. To create the DBN, the student goals were enquired after their interaction with PrimeClimb (e.g. “Have Fun”, “Avoid Failing”, “Beat Partner”, “Learn Math” and “Succeed by My Self”*)*. Conati & Maclaren (2009) are not specifying the goals, thereby giving the students excessive freedom over their learning and are not considering the nature of the educational game design, where goals can change over time. The personality theory used was Big Five (Costa & McCrae, 1992). A DBN was designed to combine the information of all the sensors and handle uncertainty.

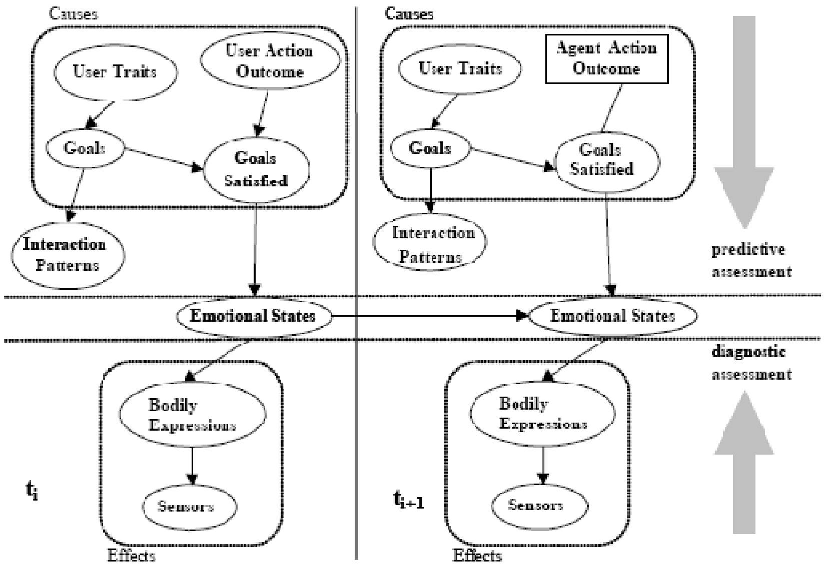

Figure 2. Affective Student Model using DBN (Conati & Maclaren, 2009, p. 4)

The EPA can provide guidance automatically or on demand. The cognitive student model was entirely separated from the affective student model. Hence, PrimeClimb uses another probabilistic model to assess the student’s knowledge about factorization. This fact is perceived as a limitation, since events originated from the student’s attempt to attain learning also have an influence over the student’s emotion. The EPA addresses the student’s needs through the provision of hints. Correlation analysis among the personality tests results, goals questionnaires and actions in the log file were used to set the probabilistic dependencies between goals, personality traits, interaction patterns and actions. Conati & Maclaren (2009) do not know how some game and agent actions will be appraised by the student against the goals, so the appraisal was based on empirical data. To label the emotional features during interaction, students’ self-reports were used, although the students that label the emotions were undergraduate students at the university and not the target population of PrimeClimb. It is noted that this fact implies limitations in the approach, since the student’s level of maturity and experience in life influences the appraisal of emotions. In addition, the approach of Conati & Maclaren (2009) cannot be used to recognise emotion over the web, since it uses hardware equipment to identify affective features of body language. Also, the participants that interact with PrimeClimb have difficulty distinguishing between the emotions originated from the game state and the emotions towards the EPA. In addition, participants were asked to register self-reports through dialog boxes, which is considered invasive. Conati & Maclaren (2009) attempted to improve their affective student model using the data collected. The evaluation was performed using the approaches of micro-average and macro-average. The former is used to recognise the percentage of cases classified correctly over all the test instances and the latter is related to the average of accurate classifications per class. Finally, Conati & Maclaren (2009) focused only on recognising four emotions with the collected data: joy, distress, admiration and reproach. The evaluation shows that the affective student model can recognise these emotions with 69.59%, 62.30%, 67.42% and 38.66% accuracy.

Easy with Eve

Easy with Eve is a VLE created by Sarrafzadeh et al. (2008) for teaching Maths at primary level. Its aim is the implementation of an ITS with the ability to recognize the student's cognitive and affective states to adapt the pedagogical strategies according to these two features. An EPA is used as output, which involves the challenges of responding with a desired personality and affective state. The EPA shows empathy and implements affinity strategies. Sarrafzadeh et al. (2008) recognised the student's affective features from facial expressions and gestures. The facial analysis was performed using a web-cam and artificial neural networks (ANNs). To classify the affective state the supervised learning method known as Supported Vector Machine (SVM) was used. Sarrafzadeh et al. (2008) tested different Kernel models, where the Radial Basis Function (RBF) method was the one that attained the best accuracy. Five-cross validation was applied to train the SVM algorithm. The approach consisted of identifying the motion of the muscles in the face by the comparison of several images per participant (Ekman & Friesen, 1978). The affective recognition system identified the affective expressions of surprise, laughter, smiling, fear, disgust and neutral with 94%, 96%, 93%, 90%, 93% and 92% accuracy. Sarrafzadeh et al. (2008) focused on recognising other gestures, since their hypothesis states that the use of gestures depicts the learner's skill level. The 2D gesture recognition system used Multilayered Feed-Forward ANNs to accomplish the analysis of 13 different types of gestures with 98.27% of accuracy. The facial and gesture recognition systems were embedded using a microcontroller, microprocessor, an Application Specific Integrated Circuit (ASIC), a Digital Signal Processor and Field Programmable Gate Arrays (FPGAs) to improve the speed and performance of the system. The embedded system is located on the server side, enabling Easy with Eve to be used via online if the students have the appropriate web-camera. The moods detected are boredom, confusion, frustration, inattention and anxiety. The accuracy of recognising these moods is not defined, since the system has not been evaluated by the final users.

The pedagogical strategies implemented were originated from the observation of videos taken while human tutors were teaching to students the concept of part-whole addition. The videos were analysed using a coding scheme to label each users interaction. Behaviours, facial expressions, their intensities and frequencies were identified in the tutor-student interaction. It is noted that human tutors adapt their strategies to the history of interaction though the tutoring session. Patterns of interaction are searched on specific learning cases. Accordingly, a weighted set of recommendations is obtained. The case-based program searches for exact matches of the specific situation. If no data is found, the query will be shortened per interaction until finding a match, but a small amount of data would be relevant to the specific scenario and relevant data will be ignored in several cases. The pedagogical strategies of Easy with Eve are giving hints, discussing problems or solutions, asking for additional information and showing an affective state according to the student's actions. The EPA uses text-to-speech generation. The case-based method used is causing problems of performance such as running out of memory. Eve cannot respond to situations, which are related to cases that are not registered in its knowledge base, therefore it is expected the implementation of an AI technique that enables Eve to learn from the learning interactions.

EMASPEL

The Emotional Multi-Agents System for Peer to peer E-learning (EMASPEL) is introduced in Neji & Ben Ammar (2007). EMASPEL was built to teach Communications Technology at undergraduate level. EMASPEL recognises the student's affective state using a web-cam and the analysis of facial features, which are classified using temporal evolution of distances. EMASPEL communicates an affective response through an emotional Embodied Conversational Agent (ECA). The EMASPEL architecture, also known as Physical conditions, Emotional state, Cognitive capabilities, and Social status (PECS) architecture, is shown in Figure 3. The *interface agent* transmits the facial information and the learner's affective state to other agents in the Multi-Agents System (MAS). Also it assigns the student's actions and information to the *agent curriculum*, the ECA, and the other agents in the MAS. PECS is comprised of three horizontal layers: the *information input layer* processes the input from the agent environment and is comprised of sensors and perception components. The *internal components layer* models the personality of the ECA and the *agent output layer* models the set of possible actions and the selection process producing the agent's response. The affective states identified by

EMASPEL are satisfaction, confidence, surprise, confusion and frustration. It is noted that in Neji & Ben Ammar (2007), the authors do not present evaluation data of the learning performance and recognition accuracy depicted by the system. EMASPEL aims to use affective feedback to bring the student to an optimal affective state to enhance learning. This ITS attempts to mirror and to assess affect in peer-to-peer (P2P) learning situations improving the student's knowledge and understanding. The *curriculum agent* saves the student's history of progress while analysing the student's profile to propose subsequent activities and sessions. The *tutoring agent* supports the learner during the learning activities and human relations, and seeks to reinforce the learner's intrinsic motivation. To enable the communication between two parts, both have to share a common knowledge of verbal and non-verbal elements (e.g. features of the current affective state). Neji & Ben Ammar (2007) made modifications to the Affective Presentation Mark-up Language (APML) introduced by DeCarolis et al. (2004) to enable the communication of a wider variety of facial expressions and moods and to enable a flexible definition of face parameters (e.g. frequency, duration and intensity) and expressions. This language received the name of EmotionStyle language, which includes distances (D_1 to D_6) to describe more accurately the student's facial expression (e.g. neutral, anger, disgust, surprise, happiness, fear and sadness). The definition of the facial distances used by EMASPEL is shown in Figure 3.

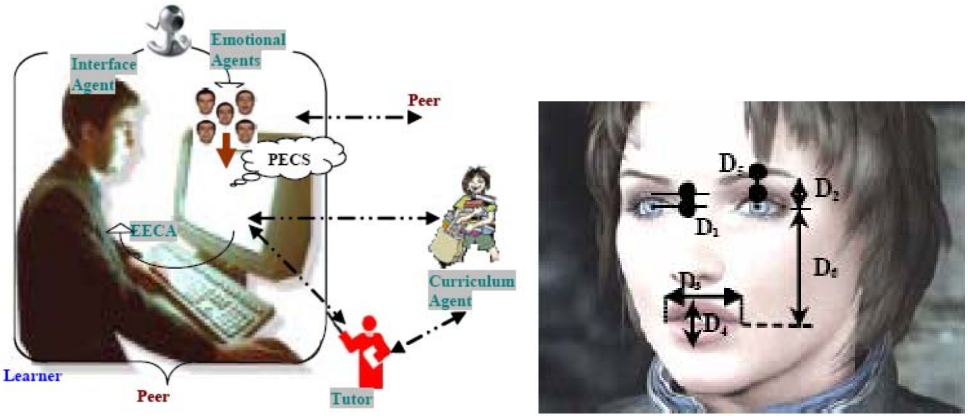

Figure 3. EMASPEL architecture (right) & Facial distances (left) (Neji & Ben Ammar, 2007, p. 127, 130)

The MAS system is organised according to the work of Ferber et al. (2003), where agents are related to groups and roles. The reasons for using a MAS with this approach when implementing an ITS are:

- A secure system can be implemented using groups with specific tasks and information, which are perceived as black boxes by other groups.
- Agents are visualised as components of a system, where dynamic modules can be added anytime.
- Semantic interoperability is guaranteed through the use of roles, which depict the obligations, requirements and skills that must be satisfied.

ESTEL

Chaffar & Frasson (2004) developed the Emotional State Towards Efficient Learning (ESTEL) system, which aimed to change the student's emotional state to one that maximises learning. ESTEL architecture is shown in Figure 4. The *Emotional Manager* is the main module that distributes and synchronises the tasks and updates the student model. The *Emotion Identifier Module* uses a sequence of colours chosen by the student to identify the student's affective state. The moods identified are happiness, disgust, anger, sadness, surprise and fear. A decision tree method was used to classify the student's selection with 57.6% accuracy. The *Personality Identifier* uses an abbreviated form of the Revised Eysenck Personality Questionnaire (Francis et al., 1992), which includes four scales: Psychoticism, Extraversion, Neuroticism and Lie. Asking to the student directly for the colour sequence that depicts the student's emotional state and asking to fill in a personality test may be perceived as intrusive and out of the learning context. A machine learning technique implemented through rules in the *Optimal Emotion Extractor* module was used to identify the optimal affective state according the student's personality. The *Emotion Inducer* was used to elicit the optimal state. Different GUIs with guided imagery, vignettes, music and images were used to induce the affective states of joy, anger, fear and sadness. This method may be perceived as out of the learning- context, since the normal flow of interaction has to be interrupted to attain the optimal affective state of the learner. The *Learner Appraiser* applies a pre-test and post-test before and after inducing the learner's optimal state, which was identified through directly enquiring 137 students about their optimal affective learning state and their personalities. Results showed that more than 28% of the students whose personality is Extraversion selected joy, 36% of the students whose personality is Lie selected confident, 29% of the students whose personality is Neuroticism selected pride and 50% of students whose personality is Psychoticism selected joy. To select the optimal affective state identified by the students, Chaffar & Frasson (2004) used a Naïve Bayes Classifier, but experimental data that validates their hypothesis and approach was not presented. Chaffar & Frasson (2004) did not indicate what the targeted domain knowledge was and did not provide detail of the pedagogical strategies used.

ERPA

Chalfoun et al. (2006) implemented the Emotional Response Predictor Agent (ERPA) that aims to infer the learner's emotional state in distant learning, which requires taking into account the student's personality traits and the events that have occurred. ERPA architecture is shown in Figure 4. The learner interacts through the GUI while the *Learner´s data acquisition component* collects data of the learner. The data related to the cognitive state is collected through quizzes and it is compared to the expected and passing scores and the quiz start time. The ID3 decision tree algorithm calculates the data gain and selects the one with the highest gain. This data is used by the *Rules Extraction component* to generate the rules and to store them in the *Rule Base*. The *emotional reaction prediction* component will use these rules to predict the student's emotional state. The learner's personality is identified through the application of an abbreviated form of the Revised Eysenck Personality Questionnaire (Francis et al., 1992). A quiz of ten questions in six different sections such Emotional Intelligence and Sports was applied to identify the student's cognitive state. At the end of the quiz the learner is asked for the expected score. In four parts of the experimental process, the expected and obtained scores are presented to the learner. The OCC model (Ortony et al., 1990) was used to select interaction events relevant to the student's emotional state. Chalfoun et al. (2006) using this approach identified disappointment, distress, joy, relief, satisfaction and fear. The decision tree also identifies the suitable cognitive training. Chalfoun et al. (2006) tested ERPA with 34 participants obtaining an accuracy of 82.4%. It is noted that Chalfoun et al. (2006) do not present results that validate this statement. ERPA can be used on-line.

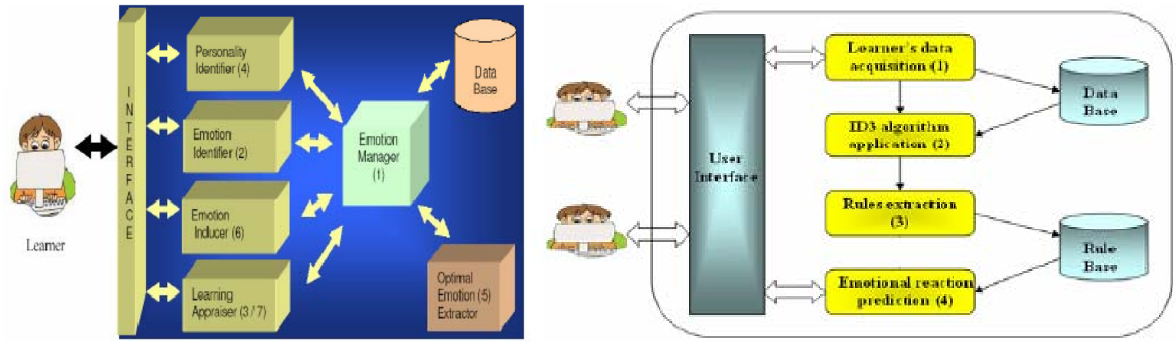

Figure 4. ESTEL (right) & ERPA (left) architectures (Chaffar & Frasson, 2004, p. 47; Chalfoun et al., 2006, p. 3)

MOCAS

Chalfoun & Frasson (2008) used a virtual reality ITS to provide affective feedback and attain adaptation to the student. A subliminal projection technique was used, which relies on the exposition of the student to information that is presented too fast to be consciously perceived and has the potential to enhance the student's learning (Watanabe et al., 2001). The subliminal technique used is *subliminal priming*, which uses a masked priming method, where the information sent for a very short period of time (e.g. 30 ms) is called a prime. This information is sent in combination with other visual information (e.g. squares and triangles) that it is not related to the information that student needs to process. The subliminal information is sent in intervals of specific time (e.g. 200 ms). The Motivational and Culturally Aware System (MOCAS), a game-like virtual learning-environment, was used to implement the subliminal priming technique (Blanchard & Frasson, 2006). MOCAS architecture, which is shown in Figure 5, is a multi-agents ITS based on the Self-determination Theory (Chirkov et al., 2003).

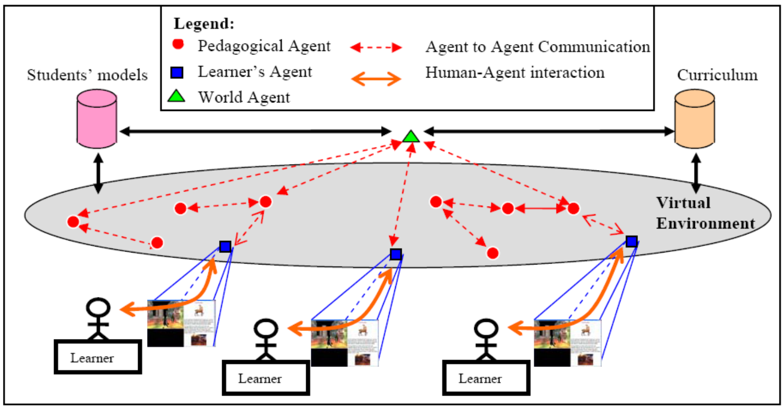

Figure 5. MOCAS architecture (Blanchard & Frasson, 2006, p. 4)

MOCAS has multiple pedagogical agents with diverse roles and attitudes. The student uses the *Learner's agent* to navigate and interact with *Pedagogical agents* in the 3D environment. The *Communication interface* is used by the student to communicate to other students and the learning content interface. The *learning content interface* adapts the domain knowledge to the student's cultural traits, is modular and enables the display of video files, questions and answers interfaces. *Multiple Pedagogical Agents* cooperate with the student's agent to provide pedagogical responses according to the teaching strategies and domain knowledge. *Multiple Learner's Agents* manage each student's interaction with the client interface, which can be modified by the learner's agent to adapt it to the student's cultural context. The *World Agent* monitors all the activity and determines the strategies of action followed by pedagogical agents (Blanchard & Frasson, 2006). It is noted that using a MAS architecture adds more challenges to the ITSs research area, since this means incorporating the problems of distributed/computer systems (e.g. ensuring the mutual exclusion over shared resources and avoiding deadlocks and livelocks issues), synchronizing and coordinating the agents activities in real-time and managing economic encounters between agents, which are concerned with their own welfare (Wooldridge, 2002). A challenge of the subliminal priming technique is displaying the subliminal information with the conscious information without degrading the performance of the system. MOCAS aims to teach the construction of an odd magic square of any order using tricks (e.g. Place the next number one square above and two squares to the right of the last one) instead of arithmetic operations or calculator. The student's must be able to deduce the trick by themselves. Chalfoun & Frasson (2008) tested MOCAS with the subliminal priming method over 31 volunteers, 16 men and 15 women all aged 28. The physiological signals (e.g. heart rate, galvanic skin response and skin temperature) from the student were acquired while interacting with MOCAS, since in the future they will be used to recognise the student's affective state. Therefore, it is noted that this method cannot easily be used in online learning. Results showed that the participant's cognitive knowledge was enhanced.

2.3 Affective and educational applications

This section reviews affective and educational applications with potential to enhance the new generation of ITSs.

2.3.1 Synthetic characters and pedagogical agents

GUIs can embody tutors or experts in other fields that use facial gestures, body language, natural language generation or playback to communicate their message. The aims are to manage the user's attention, adapt to the user through the use of AI, give affect and enhance the potential of receiving the message effectively (Johnson et al., 2000). An EPA can have two roles: learning companion (Conati & Maclaren, 2009) or tutor (Sarrafzadeh et al. 2008; D'Mello et al., 2008). EPAs and synthetic characters share the need of being believable and life-like, which is related to producing behaviour that is natural and appropriate from the user's perspective. EPAs must be able to answer the student's questions and the appearance of being knowledgeable is not enough. It is noted that this kind of response may be attained through using the knowledge representation of the domain, emotional intelligence and common sense. In the case of an EPA controlled by an ITS, this intelligent response can be attained by complementing the response with the information provided by the student model. EPAs also share with autonomous agents issues related to adapting to changes in the environment, integrating planning and execution, interacting and collaborating with other agents and achieving goals (Johnson et al., 2000). This section describes some applications of synthetic characters and EPAs that can enhance this new generation of ITSs.

This section reviews EPAs in different domains such as the Agent for Distance Learning: Light Edition (ADELE) in Shaw et al. (1999), Herman the Bug in Lester et al. (1997), Cosmo in Lester et al. (1999) and PPP Persona in André & Rist (1996). ADELE was designed for teaching medicine at undergraduate level in a VLE. ADELE architecture is shown in Figure 6 and is comprised of the *Pedagogical Agent*, which includes the *Animated-Persona module* and the *Reasoning Engine*, the *Simulation*, the *Client* and *Server* and the *Server store*. The text-to-speech generation engine is used by the *animated persona* module to produce natural language and the *Task Planner Assessor* is used by the *Reasoning Engine* to monitor and make decisions (Shaw et al., 1999). The challenges involved in this approach are the domain knowledge representation, natural language generation and animation synchronization. The knowledge representation is done in ADELE through a hierarchical plan.

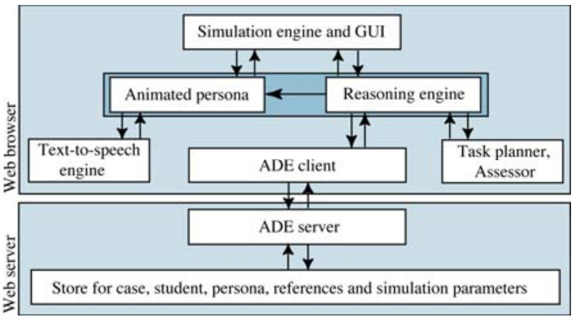

Figure 6. ADELE architecture (Shaw et al., 1999, p. 510)

Herman the Bug is the EPA of the learning environment Design-A-Plant (Lester et al., 1997), which teaches Botany to children. Herman's animated and verbal behaviours are assembled by a Real-Time Sequencing Engine (Stone & Lester, 1996), shown in Figure 7. This engine was based on the coherence-structured behaviour space approach, which includes designing a behaviour space of animation and audio segments. The segments are structured by prerequisites relationships and continuity metric. The challenge is to synchronise multiple modalities (e.g. graphics and playback) addressing context, continuity and timing. Behaviours are contextualised with problem-solving episodes and animations are contextualised with the learning environment to attain coherency (Stone & Lester, 1996). Coherency may be attained using a visual language with its own syntax and semantics such as film language and animation techniques. Stone & Lester (1996) suggest the following maxims for the design of the behaviour spaces of pedagogical agents:

- *Agent Persistence is related to maintaining an omnipresent agent onscreen to reassure the learner and increase the learner's interest.*
- *Pedagogical Object Persistence* is related to keeping the simulation model of the pedagogical task onscreen to reduce the cognitive load.
- *Agent Immersion* is related to immersing the pedagogical agent in the pedagogical task or problem by guiding its behaviour closely to the student's manipulation of the problem.
- *Verbal Support* is related to using verbalisations or verbalisations combined with little actions to offer brief reminders and interjections.

- *Contextualised musical score* is related to complementing the pedagogical agent's behaviour using context-sensitive sound tracks with appropriate tempo and instrumentation to suit the current situation.

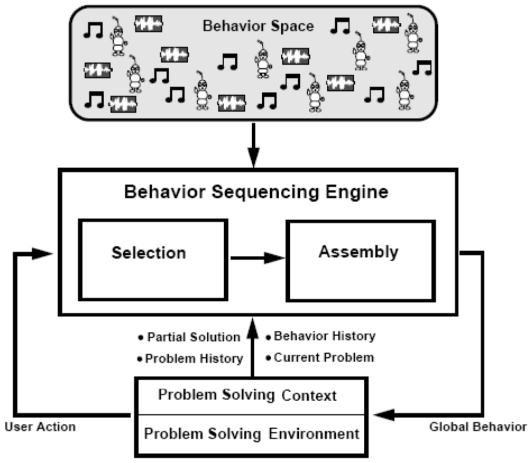

Figure 7. Herman's Behaviour Sequencing Engine (Stone & Lester, 1996, p. 428)

Behaviours have to be indexed to enable an efficient access. Stone & Lester (1996) used two kinds of ontological indexes used in explanatory behaviours, intentional indexes used in advisory behaviours and rhetorical indexes used in audio segments. The Prerequisite relations between explanatory behaviours set a partial order, where behaviours can be performed just if the behaviours on which they are dependent had been performed. The visual continuity annotations are visual attributes and features, (e.g. zoom level and frame position represented by normalised numbers) with weights that describe their priority. It is noted that indexing the media sources corresponding to the behaviours space is also a challenge.

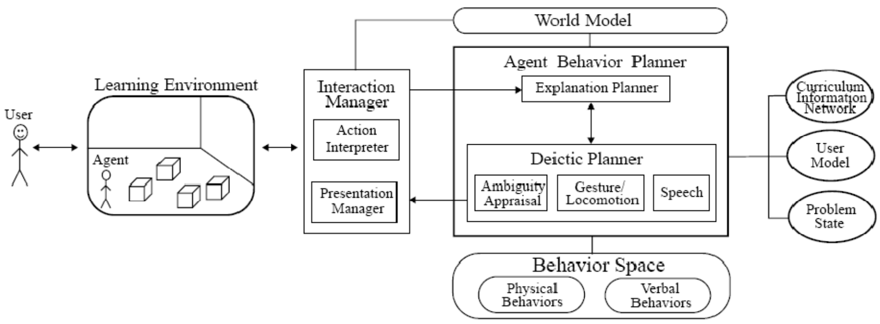

Figure 8. Cosmo architecture (Lester et al., 1999, p. 390)

Cosmo (Lester et al., 1999) is the EPA in the Internet Protocol Advisor, which was built to teach children about network routing mechanisms. This research work was focused on creating deictic believability, which is related to enabling the agent to move in the environment and to refer to the objects in it through an appropriate mixture of speech, locomotion and gestures to provide problem-solving advice. An agent behaviour planner is necessary to create deictic believability, since the features of the environment in which the agent inhabits and its history of responses must be considered to implement future responses. Deictic mechanisms should satisfy three requirements: Lack of ambiguity in their expression while communicating, immersion with the environment that surrounds the agent and support for the pedagogical intent. Cosmo architecture, known as deictic behaviour planning architecture, is shown in Figure 8. *The interaction manager* is the interface between the agent and its environment, which monitors the student's interaction to invoke the agent when the student pauses for an extended period of time or makes an error, suggesting confusion. The *explanation planner* is analogous to the discourse planner of natural language generation systems, which determines the content and structure of the sentences and plans the surface structure of the prose. The *deictic planner* receives the specifications of the *explanation planner* and makes the specifications of speech, locomotion and gestures by assigning an act, a topic, a gestural referent and spoken referent. To create the sequence of physical behaviours and verbal explanations, the *behaviour planner* analyses the knowledge representations of the world, domain and student with the current problem state and two focus stories, which correspond to gestures and speech respectively. Cosmo's behaviours are implemented in real-time using the annotations (e.g. number of frames and transition methods) made by the *behaviour planner*. There are two types of transitions: full-body behaviours and compositional behaviours. To implement non-deictic behaviours (e.g. clapping and leaning) full body images are required. In contrast, to sequence deictic behaviours (e.g. gesture and gaze) the combination of torsos, left and right arms and head are required. EPA challenges identified from this work were implementing sophisticated models of emotion, implementing conversation-based and task-oriented dialogue skills and implementing real-time natural language generation skills.

PPP persona was created by André & Rist (1996) for teaching web-based materials. PPP uses pointing gestures and synthesised speech to manage the student's attention and provide comments. Its aim was the generation of multimedia presentations in an appealing and intelligible way and enhancing human-computer communication. To overcome the challenge of identifying relations between presentation parts generated within a specific medium and parts communicated with diverse media, a directed acyclic graph (DAG) can be used to represent the rhetorical structure. The communication acts and the relations between them correspond to the nodes and the structure of the DAG. In the top of the DAG are allocated the complex communication acts and in the lowest the specifications of elementary presentation tasks. Another challenge involved in the creation of dynamic presentation is to address the temporal structure, which in the work of André & Rist (1996) is addressed through timelines positioned in a single time axis in which an event corresponds to the beginning and the end of each communication act.

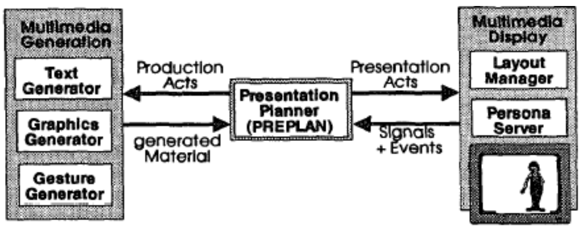

Figure 9. PPP persona architecture (André & Rist, 1996, p. 144)

The PPP persona architecture, shown in Figure 9, is a framework that defines the structure of multimedia presentations. The *presentation planner* determines the content of multimedia material (e.g. select a media combination and design a presentation script). The *medium-specific generators* inform the presentation planner when it finishes encoding the pieces of information (i.e. elementary production acts) sent by the presentation planner. The *presentation planner* uses the information sent by the generators to design the presentation script and send it to the *display components*. The *layout manager* selects effective screen layouts and maintains the user interaction. The *persona server* executes the Persona's actions assembling animation sequences. The PPP persona strategies are represented by a header (i.e. complex representation act), applicability conditions, a collection of inferior acts, a list of qualitative or metric temporal constraints and indicators of the beginning and the end of the interval. This research approach needs of presentation goals. The speech acts are sent to the *text-generator*, which determines the gesture and its coordinates. The *graphics generator* creates a window where the result will appear.

Johnson et al. (2000) identified EPA's advantages such as providing interactive demonstrations, providing navigational guidance in complex environments, providing non-verbal feedback to the learner's actions, conveying and eliciting emotion and providing adaptable pedagogical interactions. A well-designed agent can enhance the student's learning and understanding. EPAs have to be effective in helping the student to solve complex problems, since its benefits increase with problem solving complexity. It is important to note that some EPAs' architectures are similar to the architectures of ITSs, such as Cosmo's architecture. As a result, EPAs share technical challenges involved in the implementation of ITSs and interactive learning environments. In addition to these challenges Johnson et al. (2000) identified the technological issues of establishing an interface between the agent and the learning environment, designing the building blocks to generate the EPA's behaviour, developing the intelligent mechanism to adapt the pedagogical strategies and addressing the capabilities of the targeted platform or network.

SmartKom (Wolfgang et al., 2001) is a multimodal dialogue system, which uses speech, facial expressions and gestures in its input and output. Gestures and facial-expressions are acquired through video. SmartKom aims to contribute to the disambiguation of multimodal input and output on pragmatic and semantic levels. The management of interaction of SmartKom requires the use, representation and understanding of user, domain, context, task, context and media models. Smartakus, a virtual assistant, has to accomplish the user's goal by maintaining collaborative dialogue with the user and generating a plan. At end of the process, Smartakus generates and presents different results that match the user's goal. Audio and video data was segmented, translated and annotated to train machine learning methods, which can identify the user's affective state and intentions. SmartKom was implemented for diverse platforms: handheld communication, a multimodal portal and a multimodal communication kiosk. SmartKom architecture is a multi-blackboard architecture comprised of diverse modules that are classified mainly in five sets: input devices, media analysis, interaction management, application management and media design. The MultiModal Markup Language (M3L) was implemented to represent all the information that flows between modules. The information can be related to multimodal content, segmentation, synchronisation and processing results. Smartkom supports the unification and overlay operations over M3L to attain media fusion and discourse processing.

In Streit et al. (2004), Smartkom was modified to use a cognitive-model-based approach of abductive interpretation of emotions based on the OCC model. This approach consists of looking at emotions as dialogue moves that are related to explicit communication. The emotions are explained by matching the Smartakus' attitudes and actions with the current state of the situation. These attitudes and actions are compared with cognitive conditions and operators that can elicit reactions, allowing an abductive inference of explanations to the user's emotional state. Identifying the source of emotion is significant to interpreting an emotion. In SmartKom the emotion process is divided in three phases: Recognise the user's facial and prosodic affective features, collect and evaluate indications of problematic situations and the user's emotional state and interpret the user's emotional state and generate suitable responses taking into account this information. The SmartKom's *interaction module* collects and evaluates the user's indications of emotion, which are registered with a value between 0 and 1 that may change over time. The output of the *interaction module* is a set of models with a value between 0 and 1, which describe the user's specific emotions identified and features corresponding to the user's interaction. A matrix multiplication is necessary to map indicators to models. There are three main models: a model that describes the user's likelihood of being angry, a model that indicates problems during the user's interaction and a model that estimates the dialog process. Using the OCC model the conditions considered to evaluate the user's arousal are:

- The goals of the agent in the specific situation.
- The agent's actions that caused the facts in a specific situation.
- The users' goals, their attitudes to events and their standards used to judge an event.
- The facts involved with the specific situation.

Streit et al. (2004) use Abduction, an inference process to identify the best explanation, to analyse the events that may cause an emotion and to identify the one that may have the highest potential of eliciting the emotion. Two facts that influence the effectiveness of this research approach are if all the relevant explanations of an emotion were captured over time and if the intelligent system possesses the necessary skills and criteria to choose from explanations. As a result, it was noted that to evaluate the facts that can elicit an emotion, information about the user's standards, attitudes and goals is necessary. To overcome the problem of being aware of the user's goals, Streit et al. (2004) introduced meta-goals and goals. The former are related to general principles of communication and the latter are related to the user's needs. In addition, for each goal the system implements an action that can satisfy it. The validation of the identification of the user's goals was catalogued by the authors as out of the scope of the SmartKom project. For the targeted domain of SmartKom the identification process was focused on the emotion of anger. The process used to identify a possible explanation is case-based reasoning. SmartKom does not perform a verification of the user's attitudes to evaluate an event that may cause a negative emotion, since it is not considered in the SmartKom implementation. Streit et al. (2004) organised all criterion to evaluate the explanations in a decision tree and determine an appropriate response to the specific situation.

The Behaviour Expression Animation Toolkit (BEAT) was created to input written natural language and obtain a synchronised human animation (Cassell et al., 2001). BEAT aimed to overcome the challenge of synchronising non-verbal behaviour with spoken natural language in character animation. BEAT architecture is shown in Figure 10. The *language tagging module* converts the text into a parse tree adding linguistic and contextual information to schedule and assigning non-verbal behaviour. All the knowledge bases are also encoded in XML to make the process of transformation faster. The *behaviour generation module* includes a *suggestion module* and a *selection module*. The process starts with the suggestion of all the plausible behaviours and then the appropriate behaviours for specific characters are selected The *knowledge base*, the *generator set*, the *filter set* and the *translator* are data structures that can be defined by the user. The *knowledge base* represents knowledge about the world distributed corresponding to objects and actions. Each gesture is specified in a compositional notation, where shapes and trajectories of hands and arms are independently specified. The *behaviour selection module* uses a set of filters, which are applied to the tree including all the behaviour suggested. The filters delete all the behaviours that are not relevant to reflect the personality, emotional state and desirable intonation for the specific character. The *behaviour selection module* includes two types of filter strategies: conflict resolution filter (e.g. deleting all the suggestions with lower priorities) and priority threshold (e.g. diminishing the priority of the suggestion). To produce character speech two techniques can be used: Text-to-Speech (TTS) and recorded audio samples. The latter requires attaining estimates of words and phoneme timings and building an animation before its execution (e.g. Festival). In contrast, the former requires assuming the availability of real-time events while producing audio by compiling a set of rules to control the non-verbal behaviour generation (e.g. Microsoft's Whistler). BEAT can use the two approaches to create a synchronised animation.

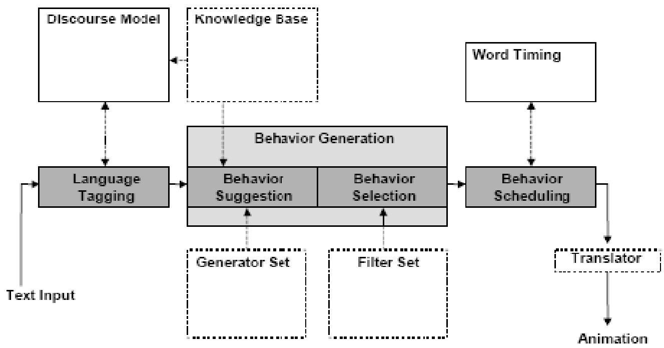

Figure 10. BEAT architecture (Cassell et al., 2001, p. 479)

The Oz Project was focused on the creation of believable agents in conjunction with interactive dramas (Mateas, 1997). Story and character play a main role in building a dramatic experience. The story signalled the goals and objectives of the interaction and the characters are comprised of physiology, psychology and sociology features. The Oz Project architecture is shown in Figure 11. It depicts a virtual world inhabited by characters. The user can interact with the world from a third person or a first person perspective. The *drama manager* monitors all the events in the world and guides the user's experience to make the history happen (e.g. changing the world model, inducing characters to pursue a course of action, adding or deleting characters). One challenge of merging storytelling and behavioural AI is to be inclined to favour the technical side and to forget the artistic goals. The Oz Project signalled that to attain character believability features such as personality, emotion, self-motivation, evolution over time, social relationships and illusion of life are necessary. In addition, the architectures of agents are suggested to attain these features, since Behavioural AI systems sense the environment, take decisions and act with an extended capability. An agent's behaviour is comprised of an action or a set of actions and it is appropriate in specific environmental and internal conditions. If the conditions suit the behaviour the agent takes action.

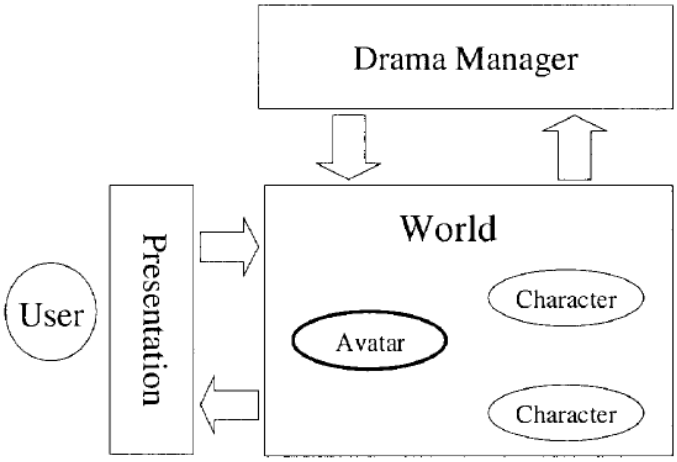

Figure 11. Oz Project architecture (Mateas, 1997, p. 298)

Hap was the language developed by the Project of Oz to specify believable agents (e.g. writing behaviours). Behaviour can be defined as a chunk of activity that can be of two types: high level (e.g. having fun) or low level (e.g. opening a door). High level goals are the character's motivators to act, which include sub-goals that must be accomplished. Each behaviour must specify the conditions that must be accomplished to enable its activation. The involved challenge is to know how many personality features must be implemented and how many degrees of freedom (DOF) are needed to implement each personality feature. The problem of this approach is the spontaneous emergence of behaviour, since the kind of behaviour that can be obtained is not ensured and it is not known until the system is executed. One of the challenges of combining story and interaction is to reach equilibrium between both, since interaction is freedom and story is predestination. The objective is to attain a user interaction that enables a structured story to happen (Mateas, 1997).

The *Drama Manager* controls the story at important moments with particular order. The moments are represented in an evaluation function. The objective is to rate the current permutations of moments against all the possible permutations of important moments in the story and to know which the possible futures are (Mateas, 1997). In addition, the *Drama Manager* monitors the user's interaction and the state of the world. A drama manager can be implemented to use the last action that has occurred or implemented to use the entire history of actions to control the story, but in the former, the user is not guided through a shaped experience. The Oz Project used a DAG to lay out scenes. The arcs of the DAG represent the must-precede relationship and until all the preceding points happened, the next plot can be included. Hints (e.g. to increase the likelihood of moving to the next scene) and obstacles (e.g. to diminish the likelihood of moving to the next scene) are used by the *Drama Manager* to influence the world and are associated with the arcs of the DAG. The granularity of control provided by the *Drama Manager* can be performed seeking to control when and where characters do their actions in the story (i.e. small grain size) or seeking to control the general direction of the story, but not the actor's actions (i.e. large grain size). The former is usually used to implement video games. In the latter, the *Drama Manager* can create a new story upon each user's interaction. The planning-problem is the main challenge of these systems.

The Oz Project implemented *A Behaviour Language* (ABL) to enable life-like characters to process several activities in parallel (e.g. gaze, speak, walk) and to enable them to react immediately. Each activity is described as a goal and each one is comprised of behaviours to accomplish it. Each sequential or parallel behaviour is comprised of goals and other behaviours, which are kept in the Active Behaviour Tree (ABT). The results corresponding to the execution of the behaviours are propagated along the ABT. Additionally the ABL allows to specify joint behaviours, which enables managing multi-agent teams. ABL was implemented to enhance and re-implement the semantic of the Hap language. The Hap agent executes a decision cycle where a leaf of the ABT is chosen to be executed. If any step of the execution of the behaviour fails, the high hierarchy behaviour fails, otherwise, it succeeds. ABL and the new modifications of Hap were used to implement the interactive-drama Façade (Mateas & Stern, 2004).

IMPROV is a system used to implement interactive actors in virtual worlds using inverse kinematics (Goldberg, 1997). In this project the scripting language is used to specify the actors' behaviour. The Behaviours have to be specified considering the internal state and the external events. To create non-deterministic characters showing different personalities and moods. The authors tune probabilities to select one action over another. However the IMPROV's scripting language does not support maintaining complex emotional states and provides more support for the animation control.

Dias & Paiva (2005) focused on creating empathic autonomous characters for teaching social education to children aged 8 to 12 using the application *Fear Not!*. *Fear Not!* enables children to watch a bullying situation performed by synthetic characters and allows them to play a part.

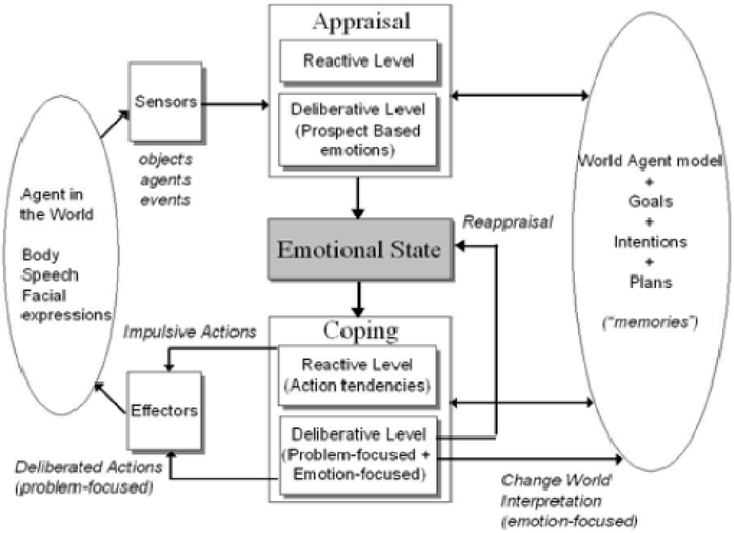

Figure 12. Fear Not! architecture (Dias & Paiva, 2005, p. 133)

The architecture of *Fear Not!* is shown in Figure 12. The architecture is comprised of two layers that enable the appraisal and coping processes. The *reactive layer* is responsible for the believable character's tendency actions and the *deliberative layer* controls the character's behaviour plans. The *continuous planner* updates all active plans (e.g. removing executed actions and changing probabilities) every time an event is received. The *deliberative level* based on the agent's plans and goals creates prospect-based emotions (Ortony et al., 1990) such as hope, fear and satisfaction. The probabilities are biased according the character's personality traits and the strategies to provide emotional response. The *reaction level* creates the fortunes-of-others emotions, well-being attraction and attribution emotions of the OCC model using a set of emotional reaction rules. The event is matched against the set of rules to generate the specific emotion. Each character pursues specific goals during the interaction, which are triggered by conditions in the environment. An activated goal is added to the partial-order plans structure. The plan with the highest priority on the structure is signalled by the identified emotion with the highest intensity. This plan will be used to continue the deliberation to choose the final response. The agent also has goals that specify protection constraints over modelling conditions that the agent wants to preserve. The planner checks all the active goals to know their likelihood of succeeding or failing.

Martinho & Paiva (2006) used an *emotivector* as a mechanism of anticipation to enhance the believability of synthetic characters. An *emotivector* is comprised of a sensor, which history of states is used to know the incoming state. Anticipation mechanisms are usually involved in the planning mechanisms of believable characters. The main advantage of an anticipation mechanism is to know if the system goes in the desired direction, which generates a reaction with an affective valence (e.g. positive or negative).

Martinho & Paiva (2006) designed and implemented the agent architecture show in Figure 13, which is used to manage multiple *emotivectors*. Each *emotivector* is related to one dimension of perception. The *emotivectors* are kept and managed by the *Salience module*. Every time a percept is sent to the processing module, the related emotivector uses the past history to compute the incoming value using a hybrid algorithm, which combines the Kalman Filter, the generalised recirculation algorithm and a model inspired in the psychology of attention originated by Posner (1980). The salience result generated by the emotivector generates an emotion according to the model inspired on the psychology of emotion created by Hammond (1970). Finally, the percept is sent to the *processing module* to provide a recommendation. *Emotivector* was tested through the implementation of a puzzle game where the synthetic flower, Aini, monitors the task progress and assists the user to solve the word puzzle by reacting to the student's actions.

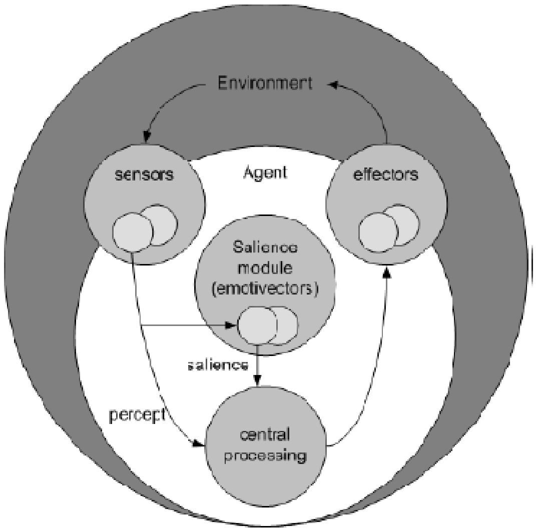

Figure 13. agent architecture (Martinho & Paiva, 2006, p. 2)

Mirage is an embodied agent inhabiting an augmented reality environment implemented by Thórrison et al. (2004) using the Constructivism Design Methodology, which is an approach that is described as a series of modules communicating through messages. This methodology offers the potential of modelling complex and multi-functional systems. In this approach, the type of message and its inputs and outputs must be specified. Each module has a role in the system. The roles can be components for perception, knowledge representation, planning, animation and other functionalities. The work of Thórrison et al. (2004) contributes to the solution of the problem of integration of broad distributed systems and enabling the collaboration between the members of the development group. One of its proposed design principles is focused on the use of blackboards with publish-subscribe functionality. The message and control flow between modules and the blackboard implementation was done using Psyclone. This design approach provides the architecture with the skills to perceive, take decisions and act.

2.3.2 Affective robots

Believable agents are implemented to act and live in engaging story worlds. By analogy, robots are built to act and move effectively in the real-world (Mateas, 1997). Therefore the creation of believable agents and robotics share a focus on embodied and situated action. In analogy with the implementation of believable characters, to attain autonomous robots that can interact with the student while learning, are necessary cognitive-affective mechanisms of attention and salience measure to define the relevance of the events, assess the value of internal or external states and guide subsequent actions and exploration. The success of these robots is measured through their ability to attain the student's knowledge, understanding and engagement (Picard et al., 2004).

iCat is a an affective robot that plays chess against a human opponent (Leite & Pereira, 2007). The player and the robot play on an electronic chess board used to monitor the player's moves. The iCat's affective state is influenced by the player's moves and the predicted final state of the interaction. Each player's move is evaluated through the application of the min-max algorithm. Then the result of the algorithm is used by an emotion model based on *emotivector* (Martinho & Paiva, 2006) to determine the iCat's affective response. The architecture of iCat is shown in Figure 14, the chess system is comprised of the *interface* with the electronic board and the *chess engine* used to evaluate the board state. The *chess system* communicates the iCat's move. The *emotion system* manages the iCat's affective state and receives the board state from the *chess system*. The *animation module* controls the blending process of animations and behaviours according to the prescription sent by the *chess system*, which manages *emotivectors* that trigger sensations that are translated into iCat's affective state.

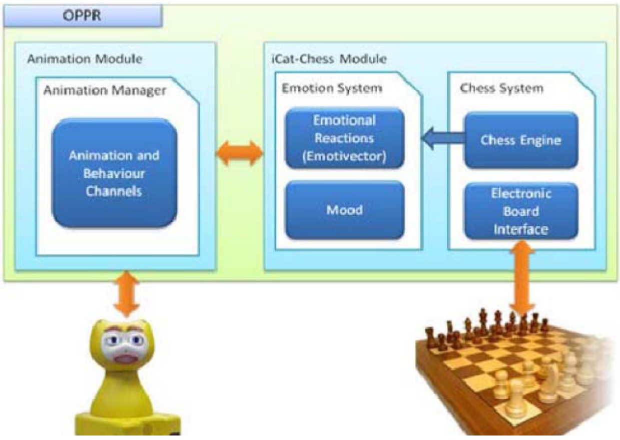

Figure 14. iCat architecture (Leite & Pereira, 2007, p. 31)

The recognition and expression of personality using computational technologies is a challenge that has been pursued using theories of personality such as Big Five (Conati & Maclaren, 2009) and Myers-Briggs (Abrahamian et al., 2004). In similar way, the recognition and expression of emotion represents a problem in the implementation of autonomous robots. Miwa et al. (2001) implemented the WE-3RV humanoid robot, which senses the user's personality and expresses personality to attain effective communication using the Big Five Theory. A 3D mental space where seven emotions (e.g. happiness, anger, disgust, fear, sadness, neutral and surprise) were mapped was defined. When a stimulus is sensed by the humanoid-robot, the stimulus is processed to determine and emotion. Then the emotion is expressed using a chosen personality from a matrix, which was used to define six personalities based on the five factors of personality and assigning intensity variables (e.g. high and low) to each factor. The 3D mental model of WE-3RV is shown in Figure 15. It is comprised of three axes: pleasantness, activation and certainty. The mental vector, M, represents the robot's mental state. Visual, auditory, tactile, temperature and olfactory sensations are processed by the mental model and the personality model.

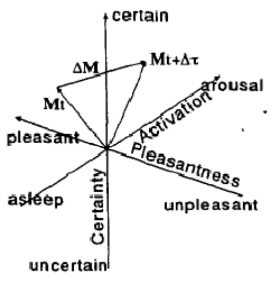

Figure 15. 3D Dynamic mental model of the WE-3RV humanoid-robot (Miwa et al., 2001, p. 1184)

2.3.3 Multimodal storytelling, game generation and affective game design

Computing and video games areas use pre-created animations or dynamic techniques (e.g. inverse kinematics) to control the animation of human motion and behaviour. To animate virtual-humans, Ma & Mc Kevitt (2006) introduced methods for blending simultaneous animations of various temporal relations. The methods were multiple channels of animation, minimal visemes for lip synchronisation and the space sites of 3D object models and virtual human, which were used to grasp and manipulate them. CONFUCIUS is a natural language visualisation and animation system (Ma & Mc Kevitt, 2006) focused on the challenges of representing temporal relations between animations and integrating animation sequences. Its architecture is shown in Figure 16. The *knowledge base* is comprised of *language knowledge* used by the *Natural Language Processing* module and *visual knowledge* used by the *animation engine* to generate animations. The *media allocator* creates the specification of the multimodal presentation in XML language and determines the content of the animation, speech and narration. The *animation engine* uses the semantic representation to create 3D VRML animations. The *synchronising module* combines the VRML animations with the speech generated by the *text-to-speech engine*. The VRML file and the presentation agent are combined by the *narration integration module*.

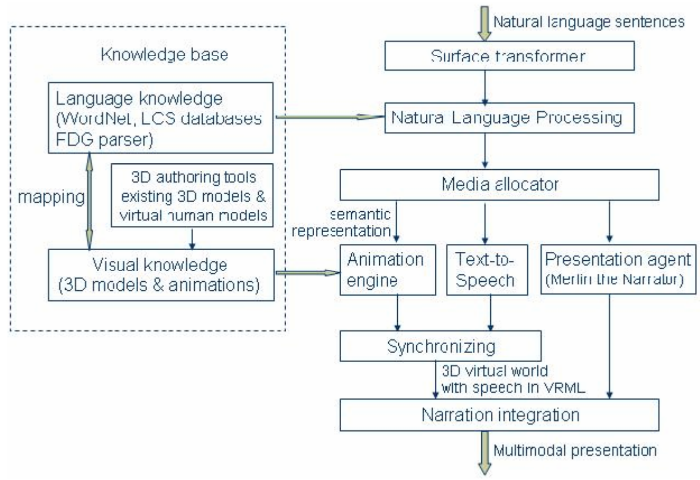

Figure 16. CONFUCIUS architecture (Ma & Mc Kevitt, 2006, p. 39)

Game generation is automatically performed by intelligent systems, which reason about the abstract rules and the rules visual realisation (Nelson & Mateas, 2007). This research area shares the problems of Creative AI involved in story and art generators, although game generation adds the problems of reasoning about dynamic, playable artefacts and allows testing game generation theories. Nelson & Mateas (2007) introduced an approach to formalising the game mechanics, which is used to generate micro-games in the style of the WarioWare series. In the approach, the authors use ConcepNet and WordNet to reason about verbs and nouns. Nelson & Mateas (2007) divided the game design problem into four interaction domains: *abstract game mechanics, concrete game representation, thematic content and control mapping*. It is important to note that the space of possible games that can be implemented depends on the knowledge of different domains contained in these systems. The *abstract game mechanics* is focused on specifying a game state and its evolution over time using knowledge about requesters, sources, requested objects, the time progression and the relationships between each of them. The *concrete game representation* is focused on representing the game state, involving the *abstract mechanics* and the player in the game-world, through an audio and visual representation. The *thematic content* is the design space involving the entire real-world and common sense knowledge used in the game, which also suggests gameplay opportunities. The *control mapping* depicts the relationship between the player's input and the effect in the game state.

Nelson & Mateas (2007) created micro-games where the user can define a game theme through specifying a verb or a noun. To attain this objective, the thematic elements of the game have to make sense and to be reasonably close, therefore a combination of ConceptNet and WordNet were used. ConceptNet is a graph-structured common-sense knowledge base comprised of the knowledge storage in the Open Mind Common Sense (OMCS) online database (Liu & Singh, 2004). The ConceptNet nodes are English words and the links represent the semantic relationships between them. Informal semantics and natural language are used in ConceptNet to easily facilitate text interaction, however ConceptNet has problems to overcome ambiguity, to respond to complex queries and its coverage is still weak. As a result, WordNet was used in combination to deal with the problem. WordNet is also a graph-structured knowledge base, although it is defined like a dictionary, where the semantic information is presented as word hierarchies. The hierarchies are used to add hierarchies to

ConceptNet. To attain similarity measures the distances between nouns in WordNet and verbs in ConceptNet are used, since WordNet has a comprehensive taxonomy and ConceptNet describes more complex notions of similarity. Nelson & Mateas (2007) defined abstract game types and types of mechanics to generate games based on the Wario Ware games. The *movement managers* match each type of game mechanics to determine the movement of the objects using the *reasoning common-sense module*, at the end of the process, the *concrete game mechanics* is comprised of the decided representation and control mappings.

Zammitto (2005) suggests that the use of colours in video games to communicate can enhance the game experience, since players use the sight sense to acquire information the 60% of the time and emotions can be increased or reduced through the use of colours. Colours have the features hue brightness and saturation. A colour can have three types of semantic meaning: innate, personal and cultural. Psychological and personality theories can involve light and colours such as the Colour Test of Lüscher (Fuentes, 2009) used to recognise personality which uses a predefined sequence of eight colours (e.g. dark-blue, blue-green, orange-red, bright-yellow, violet, brown, black and grey). The person has to select the colours from the most appealing to the least one. Lüscher's test assumes that colours have appeal or lack of appeal when associated with other colours. Lüscher's test can be used over people with colour blindness, since people react to the hue and chroma variations. Colours can be classified in two types: warm and cold colours, where the former cause the user to experience high temperature and involve a feeling of contact with the environment and the latter make the user experience low temperature and a feeling of separation. Colour brightness can cause a similar effect, since there is a positive relation between the brightness and the creation of a cold temperature perception. In addition, brighter colours can attain easily the user's attention. Colours can also be classified as saturated and unsaturated, the former are associated to a feeling of enjoyment and fun and the latter with a feeling of sadness and languidness. Colours also may be used to provide physical and psychological features to game characters. Colours can also be used in a video game to present diverse options at a glance, show the objects function or show a change in level.

2.3.4 Machines with common sense recognising affect

The area of machines with common sense was introduced by McCarthy (1959) and originated to make machines more intelligent. The final challenge is to create new interfaces, which using partial common-sense knowledge and human computer interaction (HCI) can attain their goals, assist users more effectively, attain adaptability and personalisation, create a feeling of trust and credibility and avoid mistakes and interruptions (Lieberman et al., 2004). The Problems targeted have been question-answering, story understanding and information-retrieval. These problems have been solved using approaches such as keyword matching or statistical methods. Challenges of the common sense computing area are the scarcity of common-sense knowledge bases, the system processing demands and providing a quick response. An application of common sense is the affective classification of text. Previous approaches to recognise affect in inputted text are keyword spotting, lexical affinity, statistical methods and hand-crafted models (Liu et al. 2003a). *Keyword spotting* is related to affectively classifying text and identifying affective words and their intensity. As a reference, this approach can use the affective lexicon introduced by Ortony et al. (1987), where words are divided into different categories. The two main problems of the *Keyword spotting* approach are an overreliance on surface features of the prose and also that it is not an effective method to recognise affect when a negation is involved. *Lexical affinity* is related to identifying affective words and assigning them a probability, which represents the word potential of describing a specific emotion. The limitations of the *Lexical affinity* approach are its operation at word-level and the assignation of probabilities that are biased by a particular domain. *Statistical methods* are related to use a large affective annotated text corpus as the input of a learning algorithm. The system learns the valence of the words and includes information about punctuation and word occurrence frequencies (e.g. LSA). The limitations of the *Statistical Methods* approach are its semantic weakness and the need of a large text input.

Liu et al. (2003a) used common-sense knowledge to recognise affect in inputted text. The main advantage of this approach is the recognition of affect at sentence level, which can enhance the performance of affective text-to-speech systems, context-aware systems and synthetic agents. The approach is focused on evaluating the affective features of the semantic content even when affective words are not present in the sentence structure. The approach uses the knowledge storage in the OMCS online knowledge base. The hypothesis of the approach states that there is affective knowledge involved in the attitudes to the world and everyday situations, which is combined with people's common-sense. To test the approach Liu et al. (2003a) implemented Empathy Buddy, an affective response e-mail composer. The e-mail composer is comprised of the affect sensing engine. Its process begins by processing the OMCS corpus written in a sentence-template to transforming it in binary relations. The binary relations are used by affective classification models that analyse the parsed data of the OCMS using a combination of words tagged according to the six emotions of Ekman. The affective valence of the words is propagated to the concepts related through common sense relations. This process was progressed and became ConceptNet (Liu & Singh, 2004). The text analysis process starts by segmenting the inputted text into paragraphs, sentences and independent clauses (i.e. smallest units of information that describe an event). The sentences are evaluated against the models, providing a final weighted score, which is transformed into an emotion annotation by mapping the score through disambiguation metrics. The smoother models are applied to the emotion annotation. Finally the emotion annotated sentences are expressed through an output modality. Empathy Buddy shows to the user a Chernov-style face according to the affective state depicted though the entire information inputted in the e-mail. The affect sensing engine was also used in Liu et al. (2003b) to visualise the affective structure of a web document with a multi-colour bar, where the colours correspond to each emotion. The bar is hyperlinked to have access easily to text content describing the emotion.

Li et al. (2007) presented a model to recognise emotion in text, which incorporate personality using reasoning rules. To recognise emotion the reasoning rules were built using the OCC model. Li et al. (2007) reduced the 22 emotional rules into 16 for the text features. The rules of personality were based on features of the Five-Factor model. Li et al. (2007) compared the text recognition with emotion and the text recognition with emotion and personality. Results showed that the personality factor increases the effectiveness of the emotion recognition. To implement the stated approach ConceptNet and Montylingua were used. The derived reasoning rules from the OCC model were updated according to ConceptNet. As example the reasoning rule for joy was specified as:

$$Event - consequence(txt, e) \wedge Focuson(e_1 I_1) \wedge Unexpectant(e)$$
$$\wedge\ [Desireof(I_1, e) \vee Effectof(e, f) \wedge Desireof(I_1, f)] \rightarrow joy$$

Eqn. (1) (Li et al., 2007, p. 2223)

The reasoning rule Equation (1) can be explained as: the event consequence, e, is found in the sentence and is focused on the writer, I_1, e is not expected to happen and the writer has the desire of the event, e can have the effect f and the latter is desired. If the reasoning rule is applied to the sentence "I bought a lottery ticket last week and I was very lucky to win the lottery". Using Montylingua a "win lottery" focus on self-agent is obtained. The identified events are "buy ticket" and "win lottery". Using ConceptNet, the relation $[Desireof\ 'buy\ ticket'\ 'win\ lottery']$ is found. Hence, the joy emotion is identified from processing the sentence. The authors used a vector describing the five personality dimensions and a neutral dimension. Each dimension can be assigned a value between $[-1,1]$. In addition, each dimension is classified using reasoning rules based on the personality features and the existent relations in ConceptNet. As an example Equation (2) and Equation (3) were generated to identify extroversion and introversion, which directly influence the emotions joy and distress, since extroverts enjoy being with people, are full of energy and experience positive emotions and introverts have a lack of energy, less social and quiet. Therefore, the authors decided to identify positive features for extroverts and negative features for introverts. $Synonymousspot(text, w, conversation)$ means synonymous words, w, of a conversation present in a sentence.

$$Synonymousspot(text, w, conversation) \rightarrow joy$$

Eqn. (2) (Li et al., 2007, p. 2225)

$$Synonymousspot(text, w, conversation) \rightarrow distress$$

Eqn. (3) (Li et al., 2007, p. 2225)

2.4 Knowledge representation

Knowledge representation is a sub-area of AI involved with representing related concepts, enabling a machine to attain an understanding of them (Woolf, 2009). Knowledge representation is essential to building ITSs, since their intelligent behaviour depends directly on it. In an ITS, the types of knowledge that need to be represented are knowledge about the student and the domain. These two types of knowledge must be implemented according to the pedagogical strategies. It is important to note that to develop a new generation of ITSs the type of knowledge representation influences the strategies and techniques used to recognise and provide affect. In addition, a virtual world knowledge representation is required to integrate and coordinate a multimodal output, which requires management of the semantic meaning of the media. Knowledge representation has many forms such as semantic nets, rules, constraints, plan recognition, XML documents, vectors, matrices, semantic frames and machine learning algorithms (Woolf, 2009).

Production rules in a knowledge representation language, such as MYCIN (Clancey & Buchanan, 1982) used mainly in medical informatics, but also in the implementation of ITSs (e.g. GUIDON). It has three main elements: working memory, rule base and interpreter. The working memory contains all the data acquired through solving problems. Rule base is the common data applied by the system to solve problems. On each selection-execute cycle, the interpreter is the computer module that decides which rule applies (Morgenstern, 1999). All the possible rules have to be implemented to be an effective knowledge representation. *Declarative Knowledge* is represented as text or logic statements and requires complicated procedures to enable an ITS to solve a specific problem, since the answer is deduced from facts that are found in the entire knowledge database (Woolf, 2009), such as in Carbonell (1970). It is important to note that if the knowledge representation is complex, the process of searching in the knowledge base is also complex.

Plan recognition and *machine learning techniques* use domain problems and algorithms. These techniques structure the problem into actions. Plan recognition allows the identification of the task on which the student is currently working, determining the behaviours required to update the student model and choosing the suitable pedagogical action. *Bayesian Belief Networks* (Sucar & Noguez, 2008) and *Dynamic Bayesian Belief Networks* (Reye, 1998) are often used to attain these objectives. Reye (1998) signalled that *dynamic belief networks* are necessary to reason over time about the student's knowledge evolution and the likely effects of tutorial actions. A dynamic belief network is comprised of a sequence of nodes, which represents the external state of the item and its evolution over time. To implement an ITS, *dynamic belief networks* are useful to represent changes in the student model over a sequence of interactions. The student interaction provides information about the likelihood of the student knowing the topics before and after the interaction. Reye (1998) also presented a two-phase updating algorithm of the student model. The first phase is focused on incorporating evidence about the student's knowledge state from the interaction and the second phase is focused on calculating the expected changes in the student's knowledge state after the interaction. The Two-Phase approach is illustrated in Figure 17. After the nth interaction, the probability that the student is in the state L_n depends on the fact that the student is in the state L_{n-1} and the outcome O_n. The model also includes the likelihood of the student being in the state L_n from prior interaction with the system (L_0). In addition, intelligent planning mechanisms can assists multimodal systems to choose the message that is going to be communicated, the non-verbal behaviour that will convey the message and the media that is going to be used in the output. *Influence diagrams* extend *Bayesian Belief Networks* by adding actions and utilities as nodes. The advantages are: knowing the current state, knowing the possible actions, knowing the possible results to these actions and knowing the utility of the state (Russell & Norving, 2003). *Decision Trees* are a kind of inductive learning, used also to choose the action that maximises the predicted output value of the input by taking into account a set of attributes describing the situation. The decision results are originated by retrieving values from a sequence of nodes. The aim is to improve the ability of reaction in the future (Russell & Norving, 2003). *Markov Decision Processes* are other planning mechanism used to specify a sequential decision problem using a Markov transition model (i.e. specified outcome probabilities for each action in each possible state) in a fully observable environment. The decision taken is the one that maximises the expected utility of the possible environment histories (Russell & Norving, 2003).

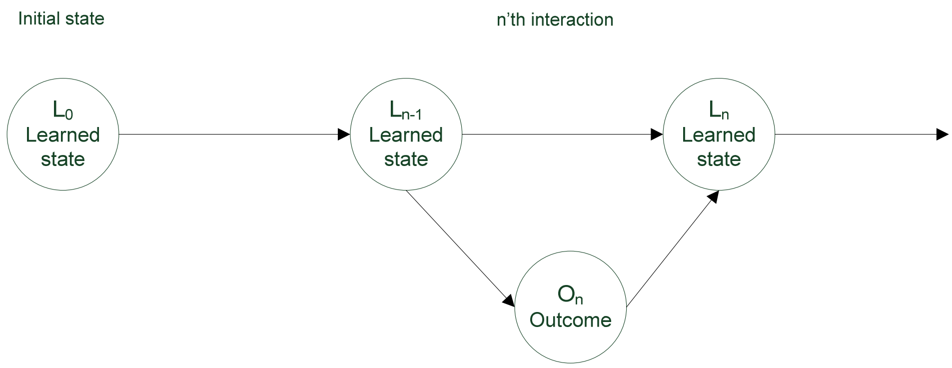

Figure 17. Two-Phase updating of dynamic belief networks (Reye, 1998, p. 276)

Vectors and *matrices* are knowledge representations used in robotics and synthetic characters to model affective and personal features (Martinho & Paiva, 2006; Miwa et al., 2001). The matrix is used to easily transform information from one domain to another. Vectors are used to know the projection of the behaviours over time. *Semantic networks* have been useful in representing relations between concepts. They are directed and undirected graphs. ConceptNet (Liu & Singh, 2004) is a knowledge base of common sense modelled as a semantic network. Its structure easily allows topic-gisting, analogy making and context-oriented inferences. The *Extensible Markup Language* (XML) was designed to easily transport and store data. Additionally, XML has been used as a form of knowledge representation. DeCarolis et al. (2004) used XML to represent the discourse plan and the communicated meanings and to create a language to generate facial expressions of the conversational agent of the MagiCster system.

Conceptual primitives and *conceptual syntax rules* were introduced in the Conceptual Dependency Theory of Schank (1972) to attain an enhanced understanding of natural language using a canonical meaning representation. The canonical representation is language-independent and assists in solving the problem of identifying the same semantic meaning as depicted by different sentences with diverse structures (e.g. John gave Mary a book and Mary received a book from John). The Conceptual primitives are atomic elements that can be combined to represent different meanings. Then when the sentences in natural language are transformed using the conceptual primitives (e.g. INGEST → to take something inside an animate object). The transformed sentences are used to apply conceptual syntax rules, attain natural language understanding and apply inference mechanisms more easily. The approach reduces the number of inferences that are stored explicitly. In addition Schank (1972) presented the use of scripts as a conceptual knowledge structure to organise memory, represent a sequence of events (i.e. tracks) related to a specific activity and to make inferences. The tracks represent many variations in the specific situation. As an example, Ma & Mc Kevitt (2006) implemented a semantic visual representation of semantic natural language based on the ideas of Schank (1972).

Semantic frames were introduced by Minsky (1974). When an event that changed the world happens and it is noted by human perception, the person selects from the mind a structured piece of information related to the event. Minsky (1974) called this chunk of memory a "frame". The piece of information is recalled from memory to be adapted according the current state of reality. According to Minsky (1974) a frame can contain information related to different aspects such as how to use it, what it is expected to happen and what to do next. The frames are comprised of terminals, which are filled in with specific data types. The terminals of each frame are pointers to the collection of the most serious problems and questions commonly associated with it. Therefore, according to Minsky (1974, p. 247) "Frames are a collection of questions to be asked about a hypothetical situation and a determined type of frame specifies the kind of issues to be raised and methods to be used in dealing with them". Hence, frames can be used to attain a visual scene analysis and to apply analogy and inference mechanisms. Brondsted et al. (1999) based on the ideas of Minsky (1974) used in CHAMELEON, a multimodal blackboard system, frames to analyse and process the multimodal input and to integrate the coordinated multimodal output.

2.5 Summary of research challenges

In sections 2.1 to 2.4, applications and research approaches corresponding to the research areas of VLEs, Educational Games, ITSs, synthetic characters and pedagogical agents, affective robots, multimodal storytelling, game generation and affective game design were reviewed. Accordingly, several research challenges were identified. The research challenges of each area are listed below.

Challenges of Intelligent Tutoring Systems (ITSs):

- Achieve an effective and flexible representation for the domain knowledge and the student model.
- Implement effective pedagogical and motivational strategies and attain equilibrium between them.
- Provide affect as response to the learner's actions.
- Identify and capture all the interaction data used to predict or recognise the student's mental state.
- Obtain the user's standards, attitudes and goals used to evaluate the interaction information.
- Recognise or predict the learner's affective state, personal traits and disposition.
- Know the degree to which ITSs must mimic human tutors.
- Identify and understand the learner's patterns and plans of interaction.
- Implement pedagogical strategies that are not perceived as intrusive and support self-development.
- Manage the uncertainty of the learner's and world knowledge data.
- Improve dialogue capabilities between the computer tutor and the student.
- Understand the semantic meaning beyond the words involved in learner-tutor communication.

Challenges of synthetic characters and pedagogical agents:

- Attain natural, coherent and appropriate believable behaviour.
- Identify and address the student's needs when answering the student's questions.
- Adapt to changes in the environment and integrate planning and execution.
- Interact and collaborate with other agents and attain the learning and artistic goals.
- Synchronise multiple modalities addressing context, continuity and timing.
- Attain an effective and flexible knowledge representation of the world, discourse, domain and the student.
- Index the media sources (e.g. animations and sounds) corresponding to the EPAs' behaviours space.
- Keep track of the history of responses, which are considered to implement future responses.
- The large size of the associated media (e.g. graphics, playback and soundtracks) databases.
- Implement sophisticated models of emotion and conversation-based and task-oriented dialogue skills.
- Implement real-time natural language generation skills to enhance the flexibility in communication.
- Implement pedagogical agents in conjunction with 3D, 2D or 2½ explanation generators.
- Identify relations between presentation parts generated within a specific medium and parts communicated with diverse media.
- Address the temporal structure of communication behaviours and acts.
- Technical challenges involved in the implementation of ITSs and interactive learning environments.
- Establish an interface between the agent and the learning environment.
- Design the building blocks to generate the EPA's behaviour.
- Develop the intelligent mechanism to adapt the pedagogical strategies.
- Address the capabilities of targeted platform or network and understand the ambiguous multimodal input.
- Relate the semantic content of verbal natural language with an underlying emotion.
- Set the prosodic features that describe the context in which speech occurs.
- Define the context information to control the intonation of the voice and hands, arms and face movements.
- Merge storytelling and behavioural AI, following technical and artistic goals.
- Know how many personality features must be implemented and how many degrees of freedom (DOF) are needed to implement each personality feature.
- Attain a user interaction that enables a structured story to happen and implement anticipation mechanisms.

Affective robots:

- Implement mechanisms of attention and salience measure to define the relevance of the events.
- Assess the value of internal or external states and plan subsequent actions and exploration.
- Recognise and show affect and personality features.

Multimodal storytelling, game generation and affective game design:

- Represent temporal relations between animations and integrate and synchronise animation sequences.
- Reason about dynamic, playable artefacts and allow the testing of game generation theories.
- Have a comprehensive knowledge of different domains to support diverse game possibilities.
- Enhance the game experience and the unambiguous communication of a message.

Machines with common sense recognising affect:

- Create interfaces that use partial common-sense knowledge and HCI to attain their goals.
- The scarcity in the common-sense knowledge bases and the system demands of processing.
- Provide a quick response and recognise affect in inputted text at sentence or event level.

3 PlayPhysics project proposal

From the stated challenges listed in section 2.5, this research project proposal is focused on solving the problems of implementing affective-motivational game based learning environments that achieve the learning goals and support the curriculum, achieving adaptability to each student, guiding and following the student performance over time using the interaction history, identifying and capturing all the interaction data involved in predicting or recognising the learner's mental state, achieving effective and flexible knowledge representations of the student, the domain and the virtual world, predicting the student's emotional state, identifying the student's personal disposition and traits, selecting and implementing motivational, affinity and pedagogical strategies, expressing affect, selecting the suitable media to communicate the message and integrating an unambiguous and synchronised multimodal output. Accordingly, in this section the research project proposal is described. This research aims to design and implement PlayPhysics, an emotional games learning environment for teaching Physics, which will be employed to evaluate the research approaches and hypotheses.

As discussed in section 2, there are two key approaches to identifying the learner's affective state in ITSs: (1) identifying the physical effects and (2) predicting emotion from its origin. PlayPhysics will employ the latter in combination with common sense to infer and ascertain the valence of the interaction events elicited by the student and the system. In addition, it was also noted that an EPA is usually employed in the game based learning environments to communicate the pedagogical, motivational or affinity response. In this regard, PlayPhysics is focused on modulating, integrating and synchronising game elements (e.g. sounds, colours and graphics) and characters to communicate an unambiguous message. To attain the stated objectives, AI techniques and planning mechanisms such as *Dynamic Bayesian Belief Networks*, *Influence Diagrams*, *Decision Trees* or *Markov Decision Processes* will be explored. PlayPhysics aims to enhance the student's motivation, learning and understanding. PlayPhysics employs the Olympia architecture as a guideline to achieve the aims and objectives of the research, therefore contributing to the state of the art of ITSs, Educational Games, Virtual Learning Environments and Affective Computing. PlayPhysics, the testbed of this research proposal, and the Olympia architecture will be evaluated in the specific case-study of teaching introductory Physics at undergraduate level. Software tools such as the Unity Game Engine and Torque Game Engine were identified and are being evaluated to accelerate the development process.

3.1 Research aims and objectives

This research proposal aims to create an emotional games learning environment, called PlayPhysics, for teaching Physics, which will incorporate a new generation ITS with the ability to predict the affective state of the learner and provide suitable and adaptable affective and pedagogical responses. Olympia architecture will be used as a guideline to implement the specific application PlayPhysics. Accordingly the research objectives are:

- Attain effective student domain knowledge, interaction events and world knowledge representations to interpret the student's actions and intentions, predict the student's affective state and identify the student's personal traits.
- Create a coordinated and adaptable multimodal output involving graphics, management of colours, sounds, game characters, an interactive story and a problem-solving approach.
- Attain effective planning, integration, synchronisation and execution mechanisms.
- Combine and provide adaptable, motivational, affinity and pedagogical strategies to enhance the student's knowledge, understanding, engagement and motivation.
- Manage the learner's interaction flow.
- Enhance Olympia architecture (Muñoz et al., 2009b) according to the aims and objectives of the research proposal.
- Design, implement and test PlayPhysics, an application using Olympia as a framework.

3.2 Rationale, hypotheses and methodology

From the literature review conducted in section 2, it was noted that predicting the learner's affective state and showing affect in response to the learner's actions can enhance the learner's motivation, knowledge and understanding. Therefore, both objectives are pursued by the new generation of ITSs. As a result, these ITSs have to manage in conjunction motivational, affinity and pedagogical strategies, which is also a challenge. In addition, it is noted that educational games are in nature highly motivational tools, although not all the educational games attain the learning goals or support learning. Educational games need to incorporate in their architectures intelligent mechanisms to follow the learner's performance and provide suitable feedback, such as ITSs. Also, games have diverse media to provide immediate feedback that can be managed to communicate an affective state. Hence, intelligent mechanisms of integrating, synchronising and planning the display of multimodal content are necessary to communicate a defined message, otherwise the information can be perceived as noise. This research aims to contribute to solve these problems through the implementation of PlayPhysics. PlayPhysics has the aim of predicting the learner's affective state. It was noted in section 2 that there are mainly two approaches to identify the learner's affective state: identifying the physical effects or predicting emotion from its origin. In addition, a combination of both can be employed. Predicting emotion from its origin involves the challenges of creating intelligent mechanisms to ascertain the affective valence of the interaction events whilst learning, knowing the likelihood of achieving the learning goals and deciding which interaction information is involved in this process. Also in section 2 it was noted that using a common sense knowledge base such as ConceptNet, the affective state can be inferred from independent clauses or the minimal description of an event, with or without the use of affective words (Liu & Singh, 2004). Based on these ideas, during the learner's interaction with the system, interaction events, which represent the learner's or system's intention to change in a specific way the virtual world to attain a specific goal, can be minimally represented through independent clauses. These independent clauses can be captured in a semantic and structured frame and later inference mechanisms can be applied to know the affective valence of the event. ConceptNet is a semantic net, in which nodes are linked through common sense relations. Each relation models uncertainty, therefore it is possible to know the negative or positive affective valence of an action with certain probability. The type of emotion originated can be determined through its cognitive origin, which is described in the OCC model (Ortony et al., 1999). The intensity of the generated emotion depends on the priorities of and the variables related to the learner's goals, standards or preferences that were affected by or related to the arousal process of the event. In addition, it is noted that by knowing the student's personality the prediction of the learner's affective state can be more accurate. Personality traits can be identified from the student's goals, the student's patterns of interaction and the selection of colours. As a result, the first hypothesis of this research proposal is that using common sense to arouse the interaction events and being aware of the learner's personal traits, the learner's affective state can be predicted over time with accuracy.

PlayPhysics needs to employ motivational, affinity and pedagogical strategies, which ensure a flow of interaction that supports the student's engagement and learning and facilitates the learner's mental state for understanding. Based on the ideas of Keller (2000), it is noted that these problems can be handled at the design stage, where by analysing the different dimensions of human motivation and the learning goals, the strategies and different media and resources can be identified, selected and combined. Then in the execution stage planning strategies based on the ideas of Dias & Paiva (2005) and Mateas & Stern (2004) can be used to define which strategy and combination of media will provide the maximum gain in student's knowledge, understanding and motivation. As a result, the second hypothesis of this research proposal is that the use of content planning mechanisms involving motivational, affinity and pedagogical strategies will provide an intelligent and synchronised multimedia output that will communicate the intention of the system unambiguously, therefore enhancing the student's understanding, knowledge and motivation. The targeted game features, which are going to be modulated, are: colours, to which semantic meaning will be mapped according to the research of Zammitto (2005) and sounds and graphics (e.g. game characters), which will be managed based on the ideas of Lester et al. (1997). The virtual world representation will be achieved based on the ideas of Ma & Mc Kevitt, (2006), Schank (1972) and Mateas & Stern (2004).

To validate the stated hypotheses, the research methodology is structured and determined as follows:

The Olympia architecture will be enhanced according to the aims and objectives of this research project proposal. Olympia was introduced in Muñoz et al. (2009b) to combine features of educational games and VLEs with ITSs. Olympia proved to be effective for enhancing the student's learning, understanding and motivation. The Olympia design will be enhanced based on the ideas of the Constructivism Design Methodology (Thórrison et al., 2004) and the web-based learner-centred paradigm (Du Boulay & Luckin, 2001). The former will be used to integrate and coordinate the different functionalities and to communicate a consistent behaviour and the latter to address the learner's needs and enable student centred learning.

PlayPhysics is an emotional games learning environment, which will be implemented using the enhanced Olympia as a guideline. The pre-analysis, analysis, design, implementation and evaluation phases of PlayPhysics will be performed using the FIDGE model (Akilli & Cagiltay, 2006) as a reference, which assists in designing, developing and implementing a game based learning environment that attains the learning goals of the curriculum. The First Principles method (Bateman & Boon, 2006) will be applied to manage a long view of the game design process. The challenges and features of the game will be designed using the principles introduced by the Games-to-teach research team (Squire et al., 2003). The PlayPhysics gameplay will be designed using DGD1 (Bateman & Boon, 2006).

The student model will be implemented using Dynamic Belief Networks (Reye, 1998) and ConceptNet (Liu & Singh, 2004). The affective student model without the learner's personality traits will be evaluated against the affective student model with the learner's personality traits using statistical methods such as linear regression. The learner's personality traits will be identified from the goals, the student's decisions and patterns of interaction based on personality theories, such as the Lüscher colour test (Fuentes, 2009) and Myers & Briggs (Bayne, 1997).

Pedagogical, motivational and affinity strategies and the tutor model will be structured according to the methodologies and theories derived from the students' opinions, expert human teachers and psychologists (Du Boulay & Luckin, 2001; Sucar & Noguez, 2008). The motivational, pedagogical and affinity strategies will be identified, selected and analysed based on the motivational design model introduced by Keller (2000). In addition, the model of Keller (2000) will be used to analyse and select the game media and features that will be employed to implement these strategies. Influence diagrams, Markov decision processes or decision trees will be evaluated for suitability in implementing the tutor model (i.e. planner) (Russell & Norving, 2003).

PlayPhysics will be used to test the approaches and hypotheses of this research proposal. Accordingly, the effectiveness of the enhanced Olympia architecture will be evaluated through the implementation of PlayPhysics, which will be employed for teaching introductory Physics to students at undergraduate level at Tecnológico de Monterrey- Mexico City Campus, Queens University Belfast and Trinity College Dublin. The students are between 17 and 23 years old. The students will be divided into control and experimental groups. Quantitative and qualitative questionnaires will be applied to the experimental group, which is comprised of students that will interact with PlayPhysics. The quantitative questionnaire will be applied to the control group to ascertain the students' knowledge and understanding of the Physics course without interacting with PlayPhysics. Experimental results will be analysed using statistical methods.

3.3 Requirements analysis

The FIDGE model suggests applying surveys to identify the learners and lecturers preferences, needs and expectations, which will be used to design a game based learning environment. Therefore, an online survey was designed and implemented. Students at undergraduate level and lecturers have participated voluntarily. Preliminary results show that the students are mainly casual game players, but there are also a large number of learners that hardly ever play a game. Also, the students prefer learning primarily by doing and secondly learning from textual and visual content. The most frequent personalities occurring amongst students, identified by applying Myers & Briggs, were ESTJ and ESTP (i.e. extrovert-sensing-thinking-judging and extrovert-sensing-thinking-perceiving). According to DGD1 (Bateman & Boon, 2006), the ESTJ personality (*Conqueror*) has a preference for the feelings of challenge and domination and the acquisition of knowledge. Challenge, strategic and puzzle games are the preferred game styles of this group of learners. The affective states of boredom, frustration and anger are predominant amongst them. Learners classified as ESTJ are goal-oriented players and the challenge must be proportional to the expected reward. The ESTP personality (*Manager*) has a preference for strategic or tactical challenge and the learners' objective is to master a process or playing methods. Learners classified as ESTP experience mainly the affective state of enjoyment for mastering a specific game. Open games (i.e. without a specific end-point) are the preferred game types for this group. From these preliminary results, it was noted that the provision of clues and payoffs, the increment in the level of difficulty, the introduction of new strategies and the embodiment of concepts are the preferred methods of feedback regarding the domain knowledge. In addition, the preferred motivational feedback was signalled as the provision of challenges, the expression of empathy, the believability of the task and the management of entertainment or humour. The online survey will be active until the end of autumn.

3.4 Olympia architecture

The Olympia architecture utilised by the PlayPhysics application is shown in Figure 18. The architecture describes a semi-open environment (Noguez & Sucar, 2006), where the learner can interact with the simulator and the virtual world to attain specific learning goals. Olympia is comprised of static and dynamic interaction modules at the GUI level. A static module does not change in real-time according to the pedagogical, motivational or affinity strategies, whilst a dynamic module is adapted by the presentation content manager module, which is part of the adaptable tutor module. *The Physics and collisions module* is comprised of the physics and maths driven objects used to enhance the level of realism in the simulation. The *Emotional feedback module* is comprised of sounds and colours that can set or change the student's mood. The *Interactive AI* module contains the Artificial Intelligence (AI) techniques used to create the believability of a learning environment. The *Input detection module* senses and handles the input. The *Networking module* controls the transmission of data across the network. The *Utilities module* is comprised of tools that assist completion of the tasks in the most efficient way. The *Scripting module* enables the external control of the application. The *Graphics rendering module* is comprised of all the graphic resources and manages the graphics and the scenes in real-time. The *Game mechanics module* manages the action-challenge relation and the game rules. The teaching and learning AI module is comprised of a new generation ITS. The *interface analysis* module identifies the interaction events that must be evaluated by the behaviour analysis module. The events are converted into frames that are kept in the short-term memory until the *behaviour analysis* module evaluates them, depending on the kind of event (e.g. cognitive or affective in nature). If the event is affective in nature, the common-sense knowledge base is used to ascertain the valence of the event (e.g. positive or negative). After the event is evaluated the evidence is propagated to the student model. Once the student model interprets the evidence, the results are sent to the tutor module to choose the affinity, motivational and pedagogical strategies that maximise the student's learning and motivation. The tutor module also has a presentation content manager module that, according to the strategies chosen by the planner, selects the media that will be used to represent and communicate them to the user. The presentation content manager module updates the world model and the game mechanics. The student model uses the history of interaction between the student and the tutor to calculate the learner's current and next mental state. The planner also uses this information to choose the most suitable strategies. Olympia architecture is *perceptual* because it predicts the student's cognitive and affective states over time, *intelligent* because it uses AI techniques (e.g. dynamic belief networks) to infer the student's mental state, *adaptable* because Olympia adjusts its pedagogical, affinity and motivational strategies to the specific situation through its intelligent planning mechanisms and *multimodal* because Olympia communicates the strategies through the modulation, integration and synchronisation of game elements and features.

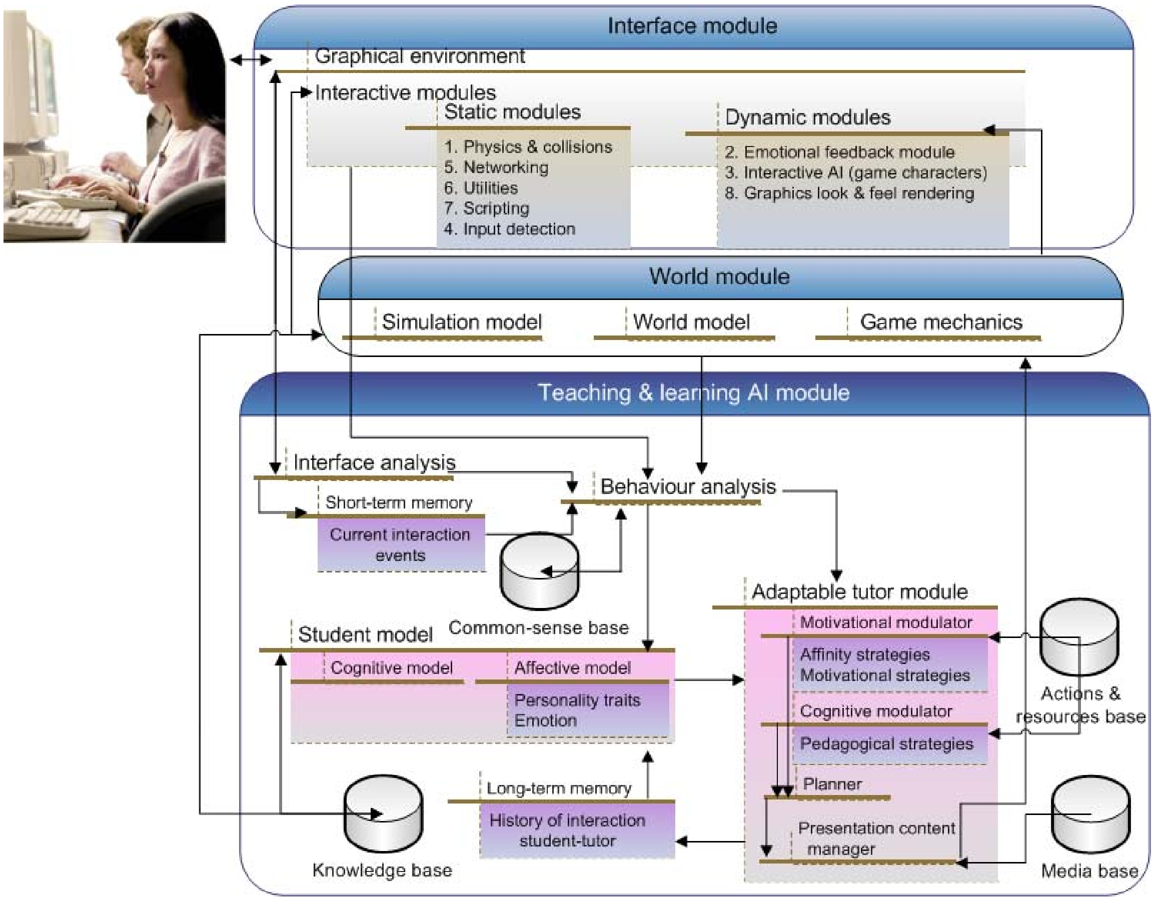

Figure 18. Olympia architecture

3.5 Comparison with previous work

PlayPhysics is compared with other new generation ITSs in Appendix A, Table A.1. These ITSs were reviewed in section 2.2.2. From this comparison, it is noted that PlayPhysics will modulate the affective features and game elements (e.g. sounds, colours, graphics, game characters) to communicate affinity, motivational and pedagogical strategies, which is the main difference with other related works that use an EPA to embody the pedagogical or affective strategies. Also, PlayPhysics will infer the affective valence of interaction events using common sense. The OCC model will be employed to ascertain the type of emotion that was elicited. In addition, PlayPhysics will use planning mechanisms to combine and select the most suitable pedagogical, motivational and affinity strategies to maximise the student's motivation, understanding and learning, which is considered a current challenge. PlayPhysics will identify the student's personality traits over time using goals, student's decisions and interaction patterns based on personality theories such as selection of colours (Fuentes, 2009) and Myers & Briggs (Bayne, 1997), which also differentiates PlayPhysics from related work, since ESTEL and ERPA used a personality questionnaire before the interaction. In addition, Prime Climb uses the learner's goals with the supposition that they do not change over time. As a result, Conati & Maclaren (2009) ascertained the students goals without considering the game and learning goals that motivated the implementation of the application in the first place. Dynamic Belief Networks, influence diagrams, Markov decision processes or decision trees will be employed to attain the intelligent features and functionalities of PlayPhysics. Accordingly, PlayPhysics' potential contributions to the state of the art of ITSs, VLEs, Educational Games and Affective Computing are:

- The interaction event and student representations that will be used to predict the learner's mental state over time.
- The provision of a coordinated, integrated and affective multimodal output involving the modulation of game features and game characters.
- Intelligent planning mechanisms that will be used to select affinity, motivational and pedagogical strategies to provide suitable feedback.
- The Olympia architecture utilised by PlayPhysics. Olympia provides a reference of coherent patterns and abstractions and is a framework to guide the system construction.
- The incorporation of all the stated elements and features in PlayPhysics, which is an integrated game learning environment that facilitates learning.

4 Software analysis

Unified Modelling Language (UML) will be applied to specify the abstract model (IBM, 2003) of PlayPhysics. Different software tools can be employed to accelerate and support the development process. ConceptNet (Liu & Singh, 2004), Torque Game Engine, Torque Game Builder (Garage Games, 2009) and Unity Game Engine (Unity Technologies, 2009), 3D Studio Max and Maya (Autodesk, 2009), Audacity (2009), Elvira (Díez, 2005) and Hugin Lite (Hugin Expert A/S, 2004), C++ (Microsoft, 2009a), C# (Microsoft, 2009b), PHP (The PHP Group, 2009) and JAVA (Sun Microsystems, 2009a), MySQL (Sun Microsystems, 2009b), Psyclone (Thórrison et al., 2004) and Haptek Automated Personalities SDK (Haptek, Inc., 2009) are software tools potentially to be used in the implementation of PlayPhysics and are being evaluated.

ConceptNet (Liu & Singh, 2004) is a common sense knowledge base and a natural language-processing tool-kit developed at MIT. It supports different context-oriented operations and inferences such as topic-gisting, affective-sensing, analogy-making and causal projection. ConceptNet is a semantic network, which is generated from data of the OMCS that is populated by online volunteers. ConceptNet can infer affect with a certain probability from independent clauses by using the stated inference mechanisms.

Torque Game Engine and Torque Game Builder are two game engines from Garage Games (2009). The former is a development platform that is supported by Mac OS X, Windows, Linux with XBox 360, Steam, i-phone and Web platforms. The Torque Script Language is a scripting language similar to C that can be employed to adapt the game behaviour. Torque Game Engine was developed in C++ and its source code can be changed by developers. The Torque Game Engine has several GUI editors and has exporters that allow the inclusion of 3D models in the game engine. The Torque Game Engine has a TorqueNet module, which supports networking. It has a Physics module that supports rigid bodies, vehicles, ragdoll, destructible objects, dynamic fluid, particle systems, destroyable joints and fluid buoyancy. Torque Game Builder is the 2D game engine of Garage Games (2009), which was simplified and generated from the 3D engine. It has diverse behaviours that can be applied to the 2D content (McGregor, 2008) and as with the Torque Game Engine, has networking, Physics and scripting modules and is targeted for the same platforms.

Unity Game Engine is a game development platform created by Unity Technologies (2009). It is integrated and extensible and can import 3D models from multiple design sources and allows the updating of them online. It has Physics and Networking modules and supports shading technologies. Unity was developed for the Mac OS X, Windows, Wii and i-phone platforms. The Scripting language of Unity is a JavaScript implementation that is faster than the script language of Flash and Director. It allows the incorporation of 3D audio and video. The games can be developed in a web browser using the Unity Web Player Plug-in. The supported browsers are Internet Explorer, Safari and Firefox. It has WWW load streamers that reduce load times. The browser colours, graphics and progress bar can be customised.

3D Studio Max and Maya are design suites of visual 3D effects, computer graphics and character animation from AutoDesk (2009). Maya incorporates fast and efficient tools, features and workflows and enables the design of rigid bodies, soft bodies, fluid dynamics, cloth, fur, hair, and new n-particles to tune shading, tools for generating plants and other natural detail, and choice of four built-in renderers, which are accessible through the new render pass architecture. In addition, Maya 2009 supports the managing of complex data without slowing the creative process through a combination of multithreading, algorithmic tuning, memory management, and tools for segmenting scenes. New animation layers and render passes enable artists to re-use existing data and minimise the iterative process. Maya 2009 also allows the optimization of tasks using Python scripting language to handle repetitive tasks and collaborative and parallel workflows. The Maya application program interface (API) was implemented in C++. Maya was targeted for the Mac OS X, Windows and Linux 64-bits platforms. 3D Studio Max allows data and scene management, the creation of parameters, organic shapes and objects to accelerate the modelling process, the subdivision of surfaces and the smoothing of polygons.

Audacity is a sound editor that also allows recording sounds. It was developed to operate in Mac OS X, Linux and Windows platforms. Cut, copy, splice or mix operations can be performed on sounds. Pitch and speed features of the recordings can be manipulated. MP3, WAV and AIFF sound files can be handled by Audacity. Audacity is open source software (SourceForge, 2009).

Elvira (Díez, 2005) is a Spanish research project, which is focused on the design and development of an environment for the implementation of algorithms (e.g. stochastic and deterministic) and methods used for probabilistic reasoning. It is a graphical interface to evaluate and create Bayesian Belief Networks and Influence Diagrams. Elvira provides canonical models (e.g. OR, AND, MAX) and decision making algorithms. Elvira can learn from databases and make a fusion of networks. Elvira was developed in JAVA and was targeted for Linux, Windows and Solaris platforms. The limitation of Elvira is the lack of help and tutorials. Hugin Lite (Hugin Expert A/S, 2004) is a decision engine, which has an API that was developed for the programming languages C, C++, .NET, JAVA and Visual Basic (e.g. Active X Server). Hugin Lite also has an intuitive GUI and a net language employed to create Bayesian Belief Networks and Influence diagrams. The net language allows specifying the structure, conditional probabilities, utility functions and the temporal order of decisions.

C++ (Microsoft, 2009a), C# (Microsoft, 2009b), PHP (The PHP Group, 2009), JAVA (Sun Microsystems, 2009a), are programming languages that are being evaluated to implement some modules of PlayPhysics providing secure access to the database, which will be defined using MySQL (Sun Microsystems, 2009b).

Psyclone (Thórrison et al., 2004) is a message-based middleware to develop distributed systems. The Psyclone publish-subscribe mechanisms routes messages to whiteboards that deliver the messages to the modules that are subscribed to them. The communication and routine protocol of Psyclone is OpenAIR.

Haptek Automated Personalities Software Development Kit (SDK) is a 3D character animation software used in web pages and standalone applications. Haptek SDK generates a character’s animation according to a given audio or speech file. It provides resources such as animations, gestures and moods, but it also includes tools to create and prototype characters. Haptek SDK incorporates APIs for ActiveX, C, C++, Internet Explorer, Netscape, Mozilla and JavaScript. Lip-synchronisation can be performed using WAV files. In addition, Haptek SDK has a compressor module that enables the compression of files to accelerate the downloading process. It also has a scripted plug-in that allows importing 3D models from 3D Studio Max. Haptek SDK provides support, documentation and tutorials.

5 Project plan

The project plan addressing the execution of this research is shown in Appendix B, Table B.1.The research plan is divided into activities, which have been classified as Milestones, Designing, testing and deploying phases, Data Collection, Experimentation, Background and focal theory, writing up thesis, Publications and Graduate Research School Training. The category *Milestones* involves due dates and key events, such as the 1st year confirmation report, the 2nd year poster presentation and 3rd year student conference presentation. Designing, testing and deploying phases are decisive activities. Data Collection includes the design and application of questionnaires to gather requirements and experimentation results. Experimentation comprises the students' interaction with the emotional games learning environment to test the research hypotheses. Background and focal theory covers the investigation performed to review the state of the art, the potential methodologies and software tools. Writing up the thesis addresses the period of time to write the doctoral thesis. A potential thesis outline is provided in Appendix C. The category Publications describes the potential targets to publish articles related to the research, such as the International Symposium of Electronic Arts (ISEA 09), the International Conference of Frontiers in Education (FIE 09) and the Journal of Transactions on Learning Technologies. Finally, the category Graduate Research School Training depicts the period in which abilities, related to the successful completion of the research project, will be attained.

According to the research project schedule, the work conducted from January to June has focused on the 1st year confirmation report. The research papers Muñoz et al. (2009a, b) were submitted and accepted for the International Symposium of Electronic Arts (ISEA 2009) and the international conference Frontiers in Education 09, respectively. Additionally, an online survey was developed to review students and lecturers about their challenges, needs and expectations while learning and teaching Physics. The survey has been applied with the objective of analysing the requirements that must be covered by the research project prototype from the users' perspective.

6 Conclusion and future work

Game based learning environments are inherently motivational, but not all of them attain the learning goals or support the curriculum. The use of design methodologies and principles and the integration of new generation ITSs to these learning environments offer a way of overcoming these challenges, although it also introduces other challenges that were covered in the literature review. These challenges include the prediction of the learner's affective state and the provision of affect in addition to selecting the most suitable pedagogical, affinity and motivational action to set the student's mental state to optimise learning and motivation. Additionally, providing a multimodal response that communicates the intention of the system unambiguously is considered a challenge. From the literature review, it was noted that ConceptNet, a common sense knowledge base, can be used to infer the affective valence of an event. As a result, the affective valence of interaction events can be inferred using common sense and their intensity can be asserted though the priority of the goals, standards and attitudes against which the event was evaluated. The type of emotion, which will be elicited, will be identified using the OCC model. In addition, from the background covered, it is noted that AI techniques such as Dynamic Belief Networks, influence diagrams, Markov decision processes and decision trees have the potential to provide a suitable representation of the student and the adaptable tutor model. Also, planning mechanisms can provide an intelligent selection of strategies that enhance the student's learning, understanding and motivation and an intelligent modulation of game features, elements and characters that can communicate the intention of the system unambiguously.

Based on the stated research ideas and in an attempt to overcome the stated challenges, this research project proposal is focused on implementing PlayPhysics, which is an emotional games learning environment for teaching Physics. PlayPhysics incorporates a new generation ITS and contributes to the state of the art of ITSs, Educational Games, VLEs and Affective Computing. Accordingly, the potential contributions of PlayPhysics are: (1) interaction events and student representations to predict the learner's mental state, (2) a coordinated, integrated and affective multimodal output involving the modulation of game features and game characters and

(3) planning mechanisms that select affinity, motivational and pedagogical strategies to provide suitable feedback. Software tools such as Haptek SDK, Torque Game Engine, Torque Game Basic and Unity Game Engine are being evaluated for the development of PlayPhysics. Olympia architecture, a reference for coherent patterns and abstractions that provides a framework to guide the system construction, is used by PlayPhysics. PlayPhysics aims to enhance the student's learning, understanding and motivation by providing suitable strategies and actions that create a high level of affinity with the learner and hence help to optimise learning performance. As future work, it is important to note that the identified methodologies to recognise emotion and personality will be tested and evaluated, as will the AI techniques used to implement the student model and the planning mechanisms of the tutor model. In addition, the definition of the PlayPhysics design and abstract model has begun.

7 References

Abrahamian, E., Weinberg, J., Grady, M. & Stanton, M.C. (2004) The effect of personality-aware Computer-Human Interfaces on Learning. *Journal of Universal Computer Science*. 4(1), 27-37.

Akilli, G. K. & Cagiltay, K. (2006) An Instructional Design Development Model for the Creation of Game-Like Learning Environments: The FIDGE Model. *In:* Pivec, M. ed. *Affective and emotional aspects of human-computer interaction*. Netherlands, Amsterdam: IOS Press, 2006, 3-7.

Alexander, S. & Sarrafzadeh, A. (2008) Foundation of an effective tutoring system: learning how human tutors adapt to student emotion. *International Journal of Intelligent Systems Technologies and Applications*. 4(3/4), 355-367.

Anderson, J. R. & Skwarecki, E. (1986) The automated tutoring of introductory computer programming. *Communications of the ACM.* 29(9), 842-849.

André, E. & Rist, T. (1996) Coping with temporal constraints in multimedia presentation planning. *Proceedings of the 13th National Conference on Artificial Intelligence , 4th – 8th August 1996, Portland, Oregon*. 1. USA: The MIT Press, 142-147.

Autodesk, Inc. (2009) *Autodesk* [online]. Available from: http://usa.autodesk.com/. [Accessed 25th May 2009].

Bateman, C. & Boon, R. (2006) *21st Century game design*. Hingham, Massachusetts: Charles River Media, Inc.

Bayne, R. (1997) The Myers-Briggs type indicator: a critical review and practical guide. 2nd ed. Cheltenham, UK: Stanley Thornes Ltd.

Bergeron, B. (2005) *Developing serious games*. Massachusetts: Charles River Media, Inc.

Bischof-Köhler, D. (2004) Empathy, Compassion and Cruelty, and How They Connect. *In: Presentation at Einstein Forum Potsdam: Zivilisationsbruch mit Zuschauer. Gestalten des Mitgefühls 9th -11th December Kaarst, Germany* [online]. Available from: http://www.bischof.com/mat/bischof-koehler_empathy.pdf. [Accessed 21st April 2009].

Blanchard, E. & Frasson, C. (2006) Easy creation of game-like virtual learning environments. *In: Workshop on teaching with robots and agents, June, 2006, Jhongli, Taiwan* [online]. Available from: http://www.iro.umontreal.ca/labs/HERON/art/blanchard_frasson_2006a.pdf . [Accessed 4th June 2009].

Brondsted, T., Dalsgaard, P., Larsen, L. B., Manthey, M., Mc Kevitt, P., Moeslund, T. B. & Olesen, K. G. (1999) Chameleon: A general Platform for Performing Intellimedia. *In:* Mc Kevitt, P. Ó Nualláin, S. & Mulvihill, C. eds. *Language, Vision and Music, Selected papers from the 8th International Workshop on the Cognitive Science of Natural Language Processing, Galway, Ireland, 1999*. Amsterdam, Philadelphia, USA: John Benjamins Publishing Company, 1999, 79-96.

Burleson, W. & Picard, R.W (2007) Evidence for Gender Specific Approaches to the Development of Emotionally Intelligent Learning Companions. *IEEE Intelligent Systems, Special issue on Intelligent Educational Systems*. 22(4), 62-69.

Carbonell, R. (1970) AI in CAI: An Artificial Intelligence Approach to Computer Aided Instruction. *IEEE Transactions on Man Machine Systems*. 11, 190-202.

Carpenter, A. & Windsor, H. (2006) *A head of the game? - Games in education, Serious Game Source* [online]. Available from: http://seriousgamessource.com/features/feature_061306_ahead_of_the_game.php. [Accessed 24th April 2009].

Cassell, J., Högni Vilhjálmsson, H. & Bickmore, T. (2001) BEAT: The Behaviour Expression Animation Toolkit. *In:* Pocock, L. ed. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques, Los Angeles, California, USA, 12th -17th August 2001*. NY, USA: ACM Press, 477-486.

Chaffar, S. & Frasson , C. (2004) Inducing Optimal Emotional state for learning in Intelligent Tutoring Systems. *In:* Goos, G., Hartmanis, J. & Van Leeuwen, J. eds. *Intelligent Tutoring Systems.* 3220. Berlin Heidelberg, Germany: Springer-Verlag, 2004, 45-54.

Chalfoun, P., Chaffar, S. & Frasson, C. (2006) Predicting the Emotional Reaction of the Learner with a Machine Learning Technique. *In: Workshop on Motivational and Affective Issues in ITS, ITS'06, International Conference on Intelligent Tutoring Systems, June, 2006, Jhongli, Taiwan* [online]. Available from: http://www-etud.iro.umontreal.ca/~chaffars/Papers/ITS%20Workshop%202006%20-%20Chalfoun%20ChaffarFrasson%20-%20FINAL.pdf.[Accessed 4th June 2009].

Chalfoun, P. & Frasson, C. (2008) Subliminal Priming Enhances Learning and Performance in a Distant 3D Virtual Intelligent Tutoring System. *IEEE Multidisciplinary Engineering Education Magazine*. 3(4), 125-130.

Chirkov, V. I., Ryan, R. M., Kim, Y. & Kaplan , U. (2003) Differentiating Autonomy from individualism and independence: A Self-determination Theory Perspective on Internalization of Cultural Orientations and Well-Being. *Journal of Psychology*. 8(1), 97-110.

Clancey, W.J. & Buchanan, B. (1982) *Exploration of teaching and problem solving strategies '1979-1982'*. Stanford, California, USA: Stanford University, (STAN-G-82-910).

Clore, G. L. & Ortony, A. (2000). Cognition in emotion: Always, sometimes, or never?. *In:* Nadel, L., Lane, R. & Ahern, G. L. eds. *The Cognitive neuroscience of emotion*. New York: Oxford University Press, 2002, 24-61.

Collins, A., Warnock, E. H., Aiello, N. & Miller, M. L. (1975) Reasoning from incomplete knowledge. *In:* Bobrow. D.G. & Collins, A. eds. *Representation and Understanding*. New York, USA: Academic Press, 1975, 383-415.

Conati, C. & Maclaren, H. (2009) Empirically Building and Evaluating a Probabilistic Model of User Affect. *User Modeling and User-Adapted Interaction* [online]. Available from: http://dx.doi.org/10.1007/s11257-009-9062-8. [Accessed 19th April 2009].

Costa, P.T, & McCrae, R.R. (1992) Four Ways Five Factors are Basic. *Personality and Individual Differences*.13, 653-665.

Czekierda, L. & Zielinski, K. (2006) Applying Workflow to Experiment Control in Virtual Laboratory. *In: Computational Science – ICCS 2006, First International Workshop on Workflow Systems in e-Science (WSES06)*. 3993. Heidelberg: Springer Berlin, 2006, 940-943.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White,H., Person, N., Kort, B., el Kaliouby, R., Picard., R.W. & Graesser, A. (2008a) AutoTutor Detects and Responds to Learners Affective and Cognitive States. *In: International Conference of Intelligent Tutoring Systems, Workshop on Emotional and Cognitive*

Issues, 23rd -27th June, Montreal, Canada [online]. Available from: http://affect.media.mit.edu/pdfs/08.dmello-etal-autotutor.pdf. [Accessed 4th June 2009].

D'Mello, S. K., Craig, S.D., Witherspoon, A., McDaniel, B.T. & Graesser, A.C. (2008b). Automatic Detection of Learner's Affect from Conversational Cues. *User modelling and User-Adapted interaction*. 18(1-2), 45-80.

DeCarolis, B., Pelachaud, C., Poggi, I. & Steedman M. (2004). APML, a mark-up language for believable behavior generation. *In:* Prendinger, H. & Ishizuka, M. eds. *Life-like Characters: Tools, Affective Functions and Applications.* Berlin Heidelberg, Germany: Springer-Verlag, 2004, 65-85.

Dias, J., Paiva, A., Vala, M., Aylett, R., Woods, S., Zoll, C. & Hall, L. (2006) Empathic characters in computer-based personal and social education. *In:* Pivec, M. ed. *Affective and emotional aspects of human-computer interaction*. Netherlands, Amsterdam: IOS Press, 2006, 246-254.

Dias, J. & Paiva, A. (2005) Feeling and Reasoning: a Computational Model for Emotional Characters. *In:* Bento, C., Cardoso, A. & Dias, G. eds. *Progress in Artificial Intelligence, 12th Portuguese Conference on Artificial Intelligence, EPIA, Covilhã, Portugal, 5th – 8th December 2005.* Heidelberg, Berlin: Springer-Verlag, 127-140.

Díez, J. (2005) *Elvira* [online]. Available from: http://www.ia.uned.es/~elvira/index-en.html. [Accessed 25th May 2009].

Du Boulay, B. & Luckin, R. (2001) Modelling human teaching tactics and strategies for tutoring systems. *International Journal of Artificial Intelligence in Education*. 12, 235-256.

Ekman, P. & Friesen, W. V. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, CA: Consulting Psychologists Press.

Evens, M. & Michael, J. (2005) *One-on-one tutoring by humans and computers*. Har/Com ed.. Philadelphia, USA: Lawrence Erlbaum Associates.

Ferber, J., Gutknecht, O. & Michael, F. (2003) From Agents to Organizations: An Organizational View of Multi-Agent Systems. *In:* Giorgini, P., Müller, J., Odell, J. eds. *Agent Oriented Software Engineering (AOSE) IV.* Berlin Heidelberg, Germany: Springer-Verlag, 2004, 443-459.

Francis, L., Brown, L. & Fhilipchalk, R. (1992) The development of an abbreviated form of the revised Eysenck Personality Questionnaire (EPQR-A). *Personality and Individual Differences*. 13, 443-449.

Freedman, R. (2000) What is an intelligent tutoring system?. *Intelligence*. 11(3), 15-16.

Fuentes, G. (2009) *Color Preferences Reveal your Personality* [online]. Available from: http://www.viewzone.com/luscher.html. [Accessed 19th May 2009].

Garage Games (2009) *Torque Game Engine* [online]. Available from: http://www.garagegames.com/. [Accessed 25th May 2009].

Ghijsen, M., Heilen, D., Nijholt, A. & op den Akker, R. (2005) Facial displays during tutoring sessions. *In: Affective Interactions: The Computer in the Affective Loop. Workshop at IUI 2005: International Conference on Intelligent User Interfaces, San Diego, CA, USA* [online]. Available from: http://eprints.eemcs.utwente.nl/1866/01/iui2005.pdf. [Accessed 4th June 2009].

Goldberg, A. (1997) IMPROV: A system for real-time animation of behaviour-based interactive synthetic actors. *In:* Trappl, R. & Petta, P. eds. *Creating Personalities for Synthetic Actors*. Berlin: Springer, 1997, 58-73.

Hammond, L. (1970) Conditioned Emotional State. *In:* Black P. ed. *Physiological Correlates of Emotion*. New York, USA: Academic Press, 1970, 245-259.

Haptek, Inc. (2009) *Haptek Automated Personalities SDK* [online]. Available from: http://www.haptek.com/developers/HaptekSDK-technical%20whitepaper.pdf . [Accessed 25th May 2009].

Heider, F. (1958) *The psychology of interpersonal relations*. New York: John Wiley & Sons.

Hsiao, H.C. (2007) A brief review of digital games and learning. *In:* Chan, T, Paiva, A., Shaffer, D. W., Kinshuk & Yang, J. eds. *The 1st IEEE International Workshop on digital game and intelligent toy enhanced learning, Jhongli, Taiwan, March 26th -28th*. Los Alamitos California: IEEE Press, 2007, 124-129.

Hugin Expert A/S (2004) *HUGIN EXPERT- Hugin Lite* [online]. Available from: http://www.hugin.com/Products_Services/Products/Demo/Lite/. [Accessed 25th May 2009].

IBM (2003) *UML basics: An introduction to the Unified Modeling Language* [online]. Available from: http://www.ibm.com/developerworks/rational/library/769.html. [Accessed 25th May 2009].

Johnson, W. L., Rickel, J. W., & Lester, J. C.(2000) Animated Pedagogical agents: Face to face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*. 11(1), 47-78.

Kincaid, J. P. (2006) *Origins and overview of modeling and simulation* [online]. Available from: http://www.peterhancock.ucf.edu/Downloads/humanfactors_2/Human%20Factors%20and%20Advanced%20Simulation%20Lecture/Overview%20of%20Modeling%20and%20Simulation.doc. [Accessed 23rd April 2009].

Keller, J. M. (2006) *What is motivational design? [manuscript]*. Florida State University, Tallahassee, Florida, USA.

Keller, J. M. (2000) Running head: Integrating motivation. How to Integrate Learner Motivation Planning into Lesson Planning: The ARCS Model Approach*. In: Presentation at VII Seminario, 23rd February 2000, Santiago, Cuba* [online]*.* Available from: http://mailer.fsu.edu/~jkeller/Articles/Keller%202000%20ARCS%20Lesson%20Planning.pdf . [Accessed 4th May 2009].

Kuo, M. J. (2007) How does an online game based learning environment promotes student's intrinsic motivation for learning natural science and how does it affect their learning outcomes?. *In:* Chan, T, Paiva, A., Shaffer, D. W., Kinshuk & Yang, J. eds. *The 1st IEEE International Workshop on digital game and intelligent toy enhanced learning, Jhongli, Taiwan, March 26th -28th*. Los Alamitos California: IEEE Press, 2007, 135-142.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998) Introduction to Latent Semantic Analysis. *Discourse Processes*. 25, 259-284.

Leite, I. & Pereira, A. (2007) iCat, THE Affective Chess Player. *In:* Petta, P. & Martinho, C. eds. *Proceedings of the Systems Demonstrations, 2nd International Conference on Affective Computing and Intelligent Interaction, 12th -14th September 2007, Lisbon, Portugal*. 4738. Heidelberg, Berlin: Springer-Verlag, 29-33.

Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B.A. & Bhogal, R. S. (1997). The persona effect: Affective impact of Animated Pedagogical Agents. *Proceeding of SIGCHI conference on Human factors in computer systems*. 359-366.

Lester, J. C., Voerman, J. L., Towns, S.G. & Callaway, C. B. (1999) Diectic believability: Coordinating gesture, locomotion and speech in life-like pedagogical agents. *Applied Artificial Intelligence*. 13, 383-414.

Li, H., Pang, N., Guo, S. & Wang, H. (2007) Research in Textual Emotion Recognition Incorporating Personality Factor. *In: IEEE International Conference on Robotics and Biomimetics, ROBIO 2007, 15th – 18th December 2007, Sanya*. USA: IEEE Press, 2222-2227.

Lieberman, H., Liu, H., Singh, P. & Barry, B. (2004) Beating Common Sense Into Interactive Applications. *AI Magazine*. 25(4), 63-76.

Liu, H., Lieberman, H. & Selker, T. (2003a) A Model of Textual Affect Sensing Using Real-World Knowledge. *In:* Johnson, L. W., Andre, E. & Domingue, J. eds. *Proceedings of the 8th International Conference on Intelligent User Interfaces, IUI 2003, 12th-15th January 2003, Miami, Florida.* USA: ACM, 125-132.

Liu, H., Selker, T. & Lieberman, H. (2003b) Visualizing the Affective Structure of a Text Document. *In:* Cockton, G. & Korhonen, P. eds. *Conference on Human Factors in Computing Systems, CHI'03, 5th- 10th April 2003, Ft. Lauderdale, Florida.* New York, NY, USA: ACM Press, 740-741.

Liu, H. & Singh, P. (2004) ConceptNet: A practical Common Sense Reasoning Toolkit. *BT Technology Journal*. 22(4), 211-226.

Liu, J., Chan, S., Hung, A. & Lee, R. (2002) Facilitators and Inhibitors of e-learning. *In:* Jain, L.C., Howlett, R.J., Ichalkaranje, N.S. & Tonfoni, G eds. *Virtual environments for teaching and learning*.1. River Edge N.J.: World Scientific, 2002, 75-107.

Ma, M. & Mc Kevitt, P. (2006) Virtual Human Animation in Natural Language Visualisation. *In:* Creaney, N. ed., *Special Issue on Research in Artificial Intelligence and Cognitive Science, Artificial Intelligence Review,* Dordrecht. (25). The Netherlands: Springer, 2006, 37-53.

Malone, T. & Lepper, M. R. (1987) Making learning fun. *In:* Snow, R. & Farr, M. eds. *Aptitude Learning and Instruction: Conative and Affective Process Analyses*. 3. Hillsdale, NJ, USA: Lawrence Erlbaum, 1987, 223-253.

Marin, B. F., Hunger, A., Werner, S. (2006) A Framework for Emotional Agents as Tutoring entities, *In:* Pivec, M. ed. *Affective and emotional aspects of human-computer interaction*. Netherlands, Amsterdam: IOS Press, 2006, 266-285.

Martinho, C. & Paiva A. (2006) Using anticipation to create believable behaviour. *In: The 21st International Conference on Artificial Intelligence and the Eighteenth Innovative Applications, 16th – 20th July 2006, Boston, Massachusetts, USA: AAAI Press* [online]. Available from: http://gaips.inesc-id.pt/gaips/shared/docs/AAAI0613MartinhoC.pdf . [Accessed 4th June 2009].

Mateas, M. (1997) *An Oz-Centric Review of Interactive Drama and Believable Agents*. Pittsburgh, PA: School of Computer Science, Carnegie Mellon University, (CMU-CS-97-156).

Mateas, M. & Stern, A. (2004) A Behaviour Language: Joint Action and Behaviour Idioms. *In:* Prendinger, H. & Ishizuka, M. eds. *Life-like Characters: Tools, Affective Functions and Applications.* Berlin Heidelberg, Germany: Springer-Verlag, 2004, 135-161.

McCarthy, J. (1959) Programs with Common Sense. *In:* Minsky, M. ed. *Semantic Information Processing*. Cambridge, MA: MIT Press, 1968, 403-418.

McFarlane, A., Sparrowhawk, A. & Heald, Y. (2001) Report on the educational use of games. *TEEM Publications* [online]. Available from: http://www.teem.org.uk/publications/teem_gamesined_full.pdf. [Accessed 24th April 2009].

McGregor, N. (2008) *Torque Engine: 3D Game Engine Review* [online]. Available from: http://videoonlinegames.suite101.com/article.cfm/torque_engine_3d_game_engine_review. [Accessed 25th May 2009]

Microsoft (2009a) *Visual C++ 2008: Express Edition* [online]. Available from: http://www.microsoft.com/Express/vc/. [Accessed 25th May 2009].

Microsoft (2009b) *Visual C# 2008: Express Edition* [online]. Available from: http://www.microsoft.com/Express/vc/. [Accessed 25th May 2009].

Minsky, M. (1974) A Framework for Representing Knowledge. *In:* Winston, P. ed. *The Psychology of Computer Vision*. New York, USA: McGraw-Hill, 1974, 211-277.

Miwa, H., Takanishi, A. & Takanobu, H. (2001) Experimental Study on Robot Personality for Humanoid Head Robot. *In: Proceedings of the International Conference of Intelligent Robots and Systems, 29th October – 3rd November 2001, Maui, Hawaii.* USA: IEEE Press, 1183-1188.

Morgenstern, L. (1999) *Knowledge Representation in Health Sciences* [online]. Available from: http://www.dbmi.columbia.edu/homepages/wandong/KR/krproductionrule.html. [Accessed 29th April 2009].

Muñoz, K., Mc Kevitt, P., Noguez, J., Lunney, T. (2009a) Combining Educational Games and Virtual Learning Environments for Teaching Physics with the Olympia Architecture. *In: International Symposium of Electronic Arts, ISEA 09, Belfast, Northern Ireland, 23rd - 30th August 2009.*

Muñoz, K., Noguez, J., Mc Kevitt, P., Neri, L., Robledo-Rella, V. & Lunney, T. (2009b) Adding Features of Educational Games for Teaching Physics. *In: The 39th IEEE International Conference Frontiers in Education, San Antonio, Texas, 18th -21st October 2009*. USA: IEEE Press.

Neji, M. & Ben Ammar, M. (2007) Agent-based Collaborative Affective e-learning framework. *The Electronic Journal of e-Learning* [online]. 5(2). Available from: http://www.ejel.org/Volume-5/v5-i2/Neji.pdf. [Accessed 8th May 2009].

Nelson, M. J. & Mateas, M. (2007) Towards Automated Game Design. *In:* Basili, R. & Pazienza, M. T. eds. *AI*IA 2007: Artificial Intelligence and Human-Oriented Computing, Proceeedings of 10th Congress of the Italian Association for Artificial Intelligence, 10th – 13th 2007, Rome, Italy*. Heidelberg, Berlin: Springer, 2007, 626-637.

Noguez, J. & Huesca, G. (2008) LaSiTo: A Lathe Simulated Virtual Laboratory. *In: Proceedings of the 38th IEEE International Conference on Frontiers in Education, 22nd -25th October 2008, Saratoga Springs, New York*. USA: IEEE Press, S2A-13-S2A-18.

Noguez, J. & Sucar, E. (2006) Intelligent virtual laboratory and project oriented learning for teaching mobile robotics. *International journal of engineering education*. 22(4), 743-757.

Norman, D.A., Ortony, A., & Russell, D.M. (2003) Affect and machine design: Lessons for the development of autonomous machines. *IBM Systems Journal*. 42, 38-44.

Oblinger, D.G. (2004) The next generation of educational engagement, *Journal of interactive media in education* [online]. (8). Available from: http://www-jime.open.ac.uk/2004/8/. [Accessed 15th April 2009].

Ohlsson, S. (1987) Some principles of intelligent tutoring. *In:* Lawler, R.W. & Yazdani, M. eds. *Learning environments and Tutoring Systems*. 1. Norwood, New Jersey, USA: Ablex Publishing, 1987, 203-237.

Ortony, A., Clore, G. L. & Collins, A. (1990) *The cognitive structure of emotions*. Cambridge, NY: Cambridge University Press.

Ortony, A., Clore, G. L. & Foss, M.A. (1987) The referential structure of the affective lexicon. *Cognitive Science*. 11, 341-364.

Pearson, N. K. & Graesser, A.C. (2008) Fourteen facts about human tutoring: food for thought for ITS developers. *CiteSeerX - Scientific Literature Digital Library and Search Engine* [online]. 6. Available from: http://www.cs.usyd.edu.au/~aied/vol6/vol6_Person.pdf. [Accessed 25th April, 2009].

Piaget, J. (1952) *The origins of intelligence*. New York, USA: International University Press.

Picard, R.W. (1995) *Affective Computing*. Mass, Cambridge: Perceptual Computing Section, Media Laboratory, Massachusetts Institute of Technology, (TR 321).

Picard, R.W., Papert, S., Bender, W., Blumberg, B., Breazeal, C., Cavallo, D., Machover, T., Resnick, M., Roy, D., & Strohecker, C. (2004) Affective learning –a manifesto. *BT Technology Journal*. 22 (4), 253-269.

Posner, M. (1980) Orienting of attention. *Quarterly Journal of Experimental Psychology*. 32(1), 3-25.

Reye, J. (1998) Two-Phase Updating of Student Models Based on Dynamic Belief Networks. *In:* Goettl, B. P., Halff, H. M., Redfield, C. I. & Shute, V. J. eds. *Proceedings of 4th International Conference of Intelligent Tutoring Systems, San Antonio Texas, USA, 16th- 19th August 1998*. 1452. Heidelberg, Germany: Springer, 174-283.

Robson, C. (1993) *Real World Research: A Resource for Social Scientist and Practitioner Researchers*. 2nd ed. Oxford, UK: Blackwell Publishers Ltd.

Russell, S. J. & Norving, P. (2003) *Artificial Intelligence: A model approach*. 2nd ed. New Jersey: Pearson Education International.

Sarrafzadeh, A., Alexander, S. Dadgostar, F. Fan ,C. & Bigdeli ,A. (2008) How do you know that I don't understand? A look at the future of intelligent tutoring systems. *Computers in Human Behavior*. 24(4), 342-1363.

Schank, R. C. (1972) Conceptual Dependency: A Theory of Natural Language Understanding. *Cognitive Psychology*. 3, 552-631.

Schank, R. C. (1994) The design of goal based scenarios. *Journal of the learning Sciences*. 3(4), 303-304.

SGI (2008) *Serious games initiative* [online]. Available from: http://www.seriousgames.org/index2.html. [Accessed 19th July 2008].

Shaw, E., Ganeshan, R., Johnson, W. L. & Millar, D. (1999) Building a case for agent assisted learning as a catalyst for curriculum reform in medical education. *In: Proceedings of the 9th International Conference on Artificial Intelligence in Education.* Amsterdam, Holland: IOS Press, 509-516.

Slator, B.M., Clark, J.T., Daniels, L.M., Hill, C., McClean, P., Saini-Eidukat, B., Schwert, D.P. & White, A.R. (2002) Use of Virtual Worlds to teach the Sciences *In:* Jain, L.C., Howlett, R.J., Ichalkaranje, N.S. & Tonfoni, G eds. *Virtual environments for teaching and learning*.1. River Edge N.J.: World Scientific, 2002, 1-40.

SourceForge (2009) *Audacity* [online]. Available from: http://audacity.sourceforge.net/. [Accessed 25th May, 2009].

Squire, K., Jenkins, H. & The Games-to-Teach Research Team (2003). Design principles of next-generation digital gaming for education. *Educational Technology*. 43(5), 17-23.

Steels, L. & Tokoro, M. (2003) *The future of learning: Issues and prospects*. Amsterdam: IOS Press.

Stone, B. A. & Lester, J. C. (1996) Dynamically sequencing an animated pedagogical agent. *Proceedings of the 13th National Conference on Artificial Intelligence, 4th – 8th August 1996, Portland, Oregon*. 1. USA: The MIT Press, 424–431.

Streit, M., Batliner, A. & Portele, T. (2004) Cognitive-Based Interpretation of Emotions in a Multimodal Dialog System. *In:* Carbonell, J. G. & Siekmann, J. eds. *Affective Dialog Systems*. 3068. Heidelberg, Berlin :Springer Verlag, 2004, 65-76.

Sucar, L.E. & Noguez, J. (2008) Chapter 10: Student Modeling. *In:* Pourret, O., Naïm, P., Marcot, B. eds. *Bayesian Networks: A Practical Guide to Applications.* United Kingdom: J. Wiley & Sons, 2008, 173-185.

Sun Microsystems (2009a) *Developer resources for Java technology* [online]. Available from: http://java.sun.com/. [Accessed 25th May 2009].

Sun Microsystems (2009b) *MySQL: The world's most popular open source database* [online]. Available from: http://www.mysql.com/. [Accessed 25th May 2009].

Sykes J. (2006) Affective gaming: Advancing the argument for game-based learning. *In:* Pivec, M. ed. *Affective and emotional aspects of human-computer interaction*. Netherlands, Amsterdam: IOS Press, 2006, 3-7.

The PHP Group (2009) *PHP* [online]. Available from: http://www.php.net/. [Accessed 25th May 2009].

Thórrison, K. R., Benko, H., Abramov, D., Andrew, A., Maskey, S. & Vaseekaran, A. (2004) Constructionist design methodology for interactive intelligences. *A.I. Magazine* [online]. 25(4). Available from: http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/1786. [Accessed 22nd May 2009].

Unity Technologies (2009) *Unity Game Engine* [online]. Available from: http://unity3d.com/unity/. [Accessed 25th May 2009].

Watanabe, T., Nanez, J. E. & Yuka, S. (2001) Perceptual learning without perception. *Nature*. 413, 5.

Wolfgang, W., Reithinger, N. & Blocher, A. (2001) SmartKom: Towards Multimodal Dialogues with Anthropomorphic Interface Agents. *In:* Wolf, G. & Klein, G. eds. *Proceedings of International Status Conference Human-Computer Interaction, DLR, Berlin, Germany, October, 23-34*[online]. Available from: http://www.smartkom.org/MTI-SmartKom.pdf. [Accessed 4th June 2009].

Wooldridge, M. (2002) *An introduction to Multi-Agent Systems*. West-Sussex, England: John Wiley & Sons Ltd.

Woolf, B. P. (2009) *Building Intelligent Interactive Tutors: Student-centered strategies for revolutionizing e-learning*. Burlington, NJ, USA: Elsevier, Inc.

Woolf, B. P. & McDonald D.D. (1984) Building a computer tutor: Design Issues. *Computer*. 17(9), 61-73.

Zammitto, V. I. (2005) The Expression of Colours. In: Changing Views: Worlds in Play, *Digital Games Research Association Digital Library* [online]. Available from: http://www.digra.org/dl/db/06278.05074.pdf . [Accessed 19th May 2009].

Appendices

Appendix A

Application	Research Reference	Tutoring modelling		Education technologies		Online	Detection of personality aspects	Feedback resource			Recognised affective features		Approach affect recognition		
		ITS[1]	IA[2]	Educational game	VLE[3]			Game modulation	Game Characters	EPA[4]	Moods	Emotions	Identifying physical effects	Predicting emotion from its origin	Using common sense and the interaction events
ESTEL	Chaffar & Frasson (2004)	✓	×	×	✓	✓	✓	×	×	×	✓	×	×	×	×
ERPA	Chalfoun et al. (2006)	✓	×	×	✓	✓	✓	×	×	×	×	✓	×	✓	×
EMASPEL	Neji & Ben Ammar (2007)	×	✓	×	✓	✓	×	×	×	✓	×	✓	✓	×	×
Easy with Eve	Sarrafzadeh et. al (2008)	✓	×	×	✓	✓	×	×	×	✓	×	✓	✓	×	×
AutoTutor	D'Mello et. al (2008a)	✓	×	×	✓	×	×	×	×	✓	✓	×	✓	×	×
MOCAS	Chalfoun & Frasson (2008)	×	✓	✓	✓	✓	×	×	×	✓	-	-	-	-	×
Prime Climb	Conati & Maclaren (2009)	✓	×	✓	✓	×	✓	×	×	✓	×	✓	✓	✓	×
PlayPhysics	Muñoz et. al (2008- 2011)	✓	×	✓	✓	✓	✓	✓	✓	×	×	✓	×	✓	✓

Table A.1. Comparison of new generation ITSs

[1] Intelligent Tutoring System
[2] Intelligent Agents
[3] Virtual Learning Environments
[4] Embodied Pedagogical Agent

Appendix B

(red)	Milestones
(orange)	Designing, testing and deploying phases
(light green)	Data collection
(dark green)	Experimentation
(blue)	Background and focal theory
(dark blue)	Writing up thesis
(gold)	Publications
(yellow)	Graduate Research School Training

Activities	2008	2009				2010				2011		
	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep
Literature Review: Background and focal theory	Background and focal theory	Background and focal theory	Background and focal theory									
Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training	Graduate Research School Training				
100 Day Review report and viva	Milestones											
39th Annual Conference, Frontiers in Education (FIE)		Publications										
Preparing a survey to find functional and non-functional requirements and justify the execution of the research		Data collection										
Applying the survey		Data collection	Data collection	Data collection	Data collection							
International Symposium of Electronic Arts (ISEA 09)		Publications	Publications									
Selecting the tools to develop the web application			Background and focal theory									
Selecting the tools to develop the 3D design of the educational game		Background and focal theory										
Selecting the tools to manage and produce sound			Background and focal theory									
Selecting the tools and methodologies to recognise aspects of emotion				Background and focal theory								
Selecting the tools and methodologies to recognise aspects of personality				Background and focal theory								
Selecting the AI tools to implement the enhanced student and adaptable tutor model					Background and focal theory							
Testing the viability of the research work by carrying a pilot study					Data collection							
Designing PlayPhysics				Designing, testing and deploying phases	Designing, testing and deploying phases							
Confirmation report and viva			Milestones									
2nd year poster presentation							Milestones					
Applying for ethical approval						Milestones	Milestones					
Developing and Implementing PlayPhysics					Milestones	Milestones	Milestones					
Testing and deploying the PlayPhysics								Designing, testing and deploying phases				
Experimentation over the subject sample								Experimentation	Experimentation			
Data collection and analysis									Data collection	Data collection		
IEEE Transactions on Learning Technologies - Journal Transactions on education								Publications	Publications			
3rd year presentation										Milestones		
Writing up thesis							Writing up thesis	Writing up thesis	Writing up thesis	Writing up thesis	Writing up thesis	Writing up thesis
Submission of thesis												Milestones

Table B.1. Research Project Schedule

Appendix C

Thesis Outline

0. Prelims. Title page, Table of Contents, List of Figures, List of Tables, Acknowledgements, Abstract, List of Acronyms and Note-on-access-to-contents.

CHAPTER-1: Introduction. This chapter provides a brief overview of the content of the Ph.D. thesis.

CHAPTER-2: Literature review. This is the literature review of previous work in area of focus of Ph.D.

CHAPTER-3: Contribution. This chapter will include the description of the unique theoretical contribution to Computer Science given in form of algorithms/heuristics.

CHAPTER-4: Olympia architecture and PlayPhysics design and implementation. This chapter will address the description of Olympia architecture and the design, implementation and development of the testbed system, PlayPhysics.

CHAPTER-5: Evaluation. This chapter will include the results of testing the Ph.D. thesis hypotheses and will provide a discussion of their evaluation using the testbed PlayPhysics.

CHAPTER-6: Conclusion. This chapter will conclude with a summary of the PhD thesis content and will provide a detailed discussion of the relation of the Ph.D. research work with other related work and future work.

7. Appendices. This section includes material such as images, code, written content and tables, which provide more insight of the research work discussed in previous sections of the Ph.D. thesis.

8. References.