Proceedings of CGAMES'2009 USA

14th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational and Serious Games

29 th July - 2 nd August 2009

Hosted by

Galt House Hotel
Louisville, Kentucky, USA

Organised by

School of Computing and Information Technology
University of Wolverhampton

And sponsored by

IEEE Computer Society Technical Committee on Simulation (TCSIM)
IEEE Computer Society (Louisville Chapter)
Society for Modelling and Simulation (SCS-Europe)
Institution of Electrical Engineers (IEE)
British Computer Society (BCS)
Digital Games Research Association (DiGRA)
Serious Games Institute, Coventry University, UK
International Journal of Intelligent Games and Simulation (IJIGS)

Edited by:

Professor Quasim Mehdi
Professor Adel Elmaghraby
Don Anderson
Dr Eugene Ch'ng

Published by The University of Wolverhampton
School of Computing and Information Technology
Printed in Wolverhampton, UK

Proceedings of CGMES'2009 Louisville, USA
Designed by Lindsey Taylor
School of Computing and Information Technology
University of Wolverhampton,
UK

Published by The University of Wolverhampton, School of Computing and Information Technology

ISBN 13-978-0-9549016-7-7

Contents

Programme Committee

General Conference Chair
Professor Quasim Mehdi
University of Wolverhampton, UK

General Programme Chair
Professor Adel Elmaghraby
University of Louisville, USA

Conference and Publicity Chair
Don Anderson
Quantum International Corporation, USA

Conference Programme Chair
Dr Eugene Ch'ng
University of Wolverhampton, UK

International Programme Committee

Professor Ian M Marshall
University of Coventry, UK

Dr. Pieter Spronck
University of Maastricht, Belgium

Dr Simon Su
Louisiana Immersive Technologies Enterprise
USA

Professor Stephane Natkin
CNAM Paris, France

Stuart Slater
University of Wolverhampton, UK

Don Anderson
Intellas Group, USA

Leon Rothkrantz
Delf University of Technology
The Netherlands

Prof. Faraon Liorens Largo
University of Alicante, Spain

Dr. Graham McAllister
University of Sussex, UK

Dr. Michel Eboueya
University of La Rochelle, France

Dr. Mariofanna Milanova
University of Arkansas, USA

Francisco José Gallego Durán
University of Alicante, Spain

Mr. Michael Callaghan
Ulster University, Belfast

Dr. Mostafa El-Said
Grand Valley State University, USA

Dr Godfried Williams
University of East London, UK

Dr Fred Mtenzi
Dublin Institute of Technology,
Ireland

Professor Colin Fyfe
University of Paisley, UK

Dr Eelke Folmer
University of Nevada, USA

Dr Gordon Calleja
IT University, Copenhagen

Dr Gareth Schott
University of Waikato, New Zealand

Mr Dan Pinchbeck
University of Portsmouth, UK

Professor Adel Elmaghraby
Louisville, University, USA

Dr. Dar-Jen Chang
University of Louisville, USA

Dr Fotis Liarokapis
The Serious Games Institute
Coventry University, UK

Don Anderson
Quantum International, USA

Dr. Rupert Simpson
University of Wolverhampton, UK

Preface

On behalf of the CGAMES conference organising committee, it is our great honour and pleasure to welcome you to the 14th International Conference on Computer Games, CGAMES 2009 Louisville, KY. CGAMES'09 USA features keynote talks and workshops on various aspects of Computer Games, which will be presented by industrial and academic speakers in the field of Serious Games and Virtual Worlds. This year, a special session on Microsoft XNA development will accompany the talks. The conference is co-sponsored by the British Computer Society (BCS), IEEE Computer Society Louisville Chapter, IEEE Louisville Section, IEEE Computer Society Technical Committee on Simulation (TCSIM), University of Wolverhampton, University of Louisville, USA, and the Intellas Group, USA. We are very privileged to be able to bring the conference to Louisville, KY USA.

The theme has been chosen to reflect the major changes in the way in which games are being developed and played. Over the last decade, CGAMES has served the computer games community by promoting the advancement, innovation, and potential applications of computer games to other areas such as training, education, social studies, and health. The 13th conference was held at the Light House, Wolverhampton, UK, from 3rd - 5th November 2008, which attracted a significant audience from both the games industry and academia. At present the CGAMES conference continues to maintain constructive links with: *The IEEE Computer Society*, *British Computer Society*, *Digital Games Research Association (DiGRA)*, and the *Society for Modelling and Simulation (SCS)*. We would like to express our gratitude toward these organisations for their invaluable assistance and support that make the conference a success.

The conference aims to promote both established and new researchers, including postgraduate research students who may be presenting their work for the first time. The quality of submitted papers continues to improve each year and our special thanks go to the reviewers who have been most diligent in their responses by providing detailed and informative feedback to authors. We are very pleased to have IEEE TCSIM sponsoring the best paper award at CGAMES 09 USA.

We would like to take this opportunity to announce The 15th International conference: Artificial Intelligence, Animation, Mobile Systems, Educational and Serious Games, which will be hosted by The University of Alicante, Spain on 2nd - 4th November 2009.

The Conference Programme Committee would like to express their appreciation for the hard work put in by our International Programme Committee members. We are grateful or their generosity, time, and valuable support given for organising the conference. This conference is continually growing thanks to the hard work put in by the International Programme Committee members. We fully appreciate their generosity, time, and effort to help organise this conference.

We trust that you will all enjoy your stay in Louisville, the home of Churchill Downs and the Kentucky Derby, and benefit from this conference by making new contacts for future mutual collaboration.

Professor Quasim Mehdi, On behalf of the Conference Programme Committee
University of Wolverhampton, July 2009

Proceedings

On the Future of Serious Games in Science and Industry*

Dr. L. Miguel Encarnação
Humana Inc. - Innovation Center
500 W Main St.
Louisville, KY 40202, USA
Email: lme@computer.org

KEYWORDS

Serious Games, taxonomy, persuasive technology, business, applications, research, challenges.

ABSTRACT

The availability of high-performance commodity graphics hardware at low cost allowed the convergence of two historically distinct application areas of Computer Graphics – video games and interactive visual simulation. The result is the newly emerging R&D area of Serious Games at the nexus of videogame design, pedagogy, human-computer interaction, simulation, AI and digital storytelling. It is aimed at exploring the opportunities and challenges of using entertaining game play for real-world applications beyond mere entertainment towards for instance education, training, motivation, rehabilitation, persuasion and decision making.

True success of Serious Games depends on consumers' realization of the inherent value of those products for various self-efficacy applications in life. Additionally, there exist numerous challenges for Serious Games developers including smaller consumer audiences, higher development cost, the need for embedded evaluation and assessment functionality, as well as insufficient business models and distribution channels. Those challenges make it traditionally hard to compete against popular entertainment titles in the consumer retail market. It might, however, well be this additional value that has the potential to give Serious Games – if well designed and properly targeted – a competitive edge over mere entertainment titles in the current recession-plagued industry.

This paper is providing the rationale behind the corresponding keynote presentation and aims at giving an overview on the field of Serious Games from an application industry perspective. It further aims at discussing and identifying requirements for Serious Games R&D towards effectiveness as well as consumer acceptance and adoption.

INTRODUCTION – ENTERTAINMENT GAME MARKET UNDER SIEGE

The jury is still out on whether the current slowdown in hardware and software sales in the gaming industry is just a well-deserved rest after an impressive sprint in the fall of 2008 or a sign that the world-wide recession makes consumers focus more on life essentials than on dispensable luxury items like game consoles and entertainment. However, a closer look at recent developments in the gaming industry seems to support the hypothesis that pure entertainment titles are being hit hard in the current economy: In May of 2009, Nintendo forecast an 11.8 percent decline in operating profit in the year ending March 31, 2010. For the fourth fiscal quarter ending March 31, 2009, Nintendo's sales dropped 15 percent from a year earlier, and its net income was off 43 percent from a year ago. While this outlook reflects a slowdown, Nintendo is still far better off than the game divisions of Sony and Microsoft. In May video game sales fell by 23 percent from the same period a year ago, according to the NPD Group, which tracks domestic sales. (NPD 2009) Sales of games hardware (consoles and handheld devices) dropped by 30 percent, while sales of software fell by 17 percent. As a market analyst states: "This, the third consecutive month of declining year-over-year sales, leaves scant support for those arguing that the stay-at-home-and-play industry is not susceptible to recessionary pressures".

Additional news on layoffs in Microsoft's gaming division, EA Sports, and the bankruptcy news and discussions around Midway Games (*Mortal Combat*), Deadline Games (*Watchmen*), and Interplay (*Fallout*) are doing their part in establishing the perception of an industry-wide down-turn.

However, the continuing success of the Nintendo Wii technology and market news about major non-obvious partnerships in the entertainment games market, seem to highlight a trend towards finally exploring more purposeful uses of game play in order to increase access to niche markets.

SERIOUS GAMES TO THE RESCUE

In 2007, Eliane Alhadeff – a recognized Serious Games blogger and market observer (Alhadeff 2007) –

* This paper is based on the keynote presentation by the author to the 14th International Conference on Computer Games titled "Recession-proof Game Development – Putting the aim in Games."

predicted the Serious Games market to grow to 2 billion dollars in revenues, initially based strongly on corporate training, healthcare simulation, and learning games, who she predicted to become mainstream by 2012. Whether these predictions still hold up in the current recession and with the increasing cost for game development is not clear, however recent announcements at least indicate that the entertainment market has taken notice of the growing potential of Serious Games as well as the increasing interest of casual gamers in games with a purpose. What the Wii success taught us is that in addition to traditional hardcore gamers, casual gamers are forcefully entering the game consumer market and those do not see their purpose in life sitting all day inactively in front of a console. Across all demographics, those casual gamers don't and might never see themselves as gamers and are attracted by activities that add additional value to the mental, physical, intellectual or social well-being. Furthermore, due to the social component of this game play – 59 percent of gamers play with other gamers in person – there is a unification happening between hardcore and casual gamers.

Consequently, a new type of game is evolving, which aims at seducing the new casual gamer generations through active, social, cognitive or educational play without requiring them to pass through a steep learning curve during seemingly endless hours of game play. Products like Wii, WiiFIT, eyeToy, Konami's DDR, and Microsoft Natal on the hardware and peripheral side, and WiiSports, Guitar Hero, RockBand, and EA Active's Virtual Trainer on the software side are clearly indicating a new direction for the entertainment games market.

Furthermore, EA's partnership with toy manufacturer Hasbro to launch family classics, Hasbro's competitor Mattel closing an exclusivity deal with Neuorsky the leader in consumer brain computer technologies (towards launching a new category of brain games and toys that operate using the power of concentration), Jerry Bruckheimer's opening of a games studio in partnership with MTV games, Warner studios' purchase of midway games; all those developments show the convergence of the traditional entertainment video games markets with other entertainment genres, channels, and leisure offerings in response to having to serve a broader, more diverse, and more demanding gamer audience. Gaming is becoming a ubiquitous part of life and is in this sense providing new opportunities for Serious Games to evolve and being takes seriously.

WHAT'S IN A GAME, SERIOUSLY?!

Looking at the whole space currently referred to as Serious Games, it is clear that this term is being attributed to many different aspects of video game

Table 1: Preliminary Serious Games taxonomy (Sawyer and Smith 2008)

	Games for Health	Advergames	Games for Training	Games for Education	Games for Science and Research	Production	Games as Work
Government & NGO	Public Health Education & Mass Casualty Response	Political Games	Employee Training	Inform Public	Data Collection / Planning	Strategic & Policy Planning	Public Diplomacy, Opinion Research
Defense	Rehabilitation & Wellness	Recruitment & Propaganda	Soldier/Support Training	School House Education	Wargames / planning	War planning & weapons research	Command & Control
Healthcare	Cybertherapy / Exergaming	Public Health Policy & Social Awareness Campaigns	Training Games for Health Professionals	Games for Patient Education and Disease Management	Visualization & Epidemiology	Biotech manufacturing & design	Public Health Response Planning & Logistics
Marketing & Communications	Advertising Treatment	Advertising, marketing with games, product placement	Product Use	Product Information	Opinion Research	Machinima	Opinion Research
Education	Inform about diseases/risks	Social Issue Games	Train teachers / Train workforce skills	Learning	Computer Science & Recruitment	P2P Learning Constructivism Documentary?	Teaching Distance Learning
Corporate	Employee Health Information & Wellness	Customer Education & Awareness	Employee Training	Continuing Education & Certification	Advertising / visualization	Strategic Planning	Command & Control
Industry	Occupational Safety	Sales & Recruitment	Employee Training	Workforce Education	Process Optimization Simulation	Nano/Bio-tech Design	Command & Control

design, development, hardware and game play. Maybe this is just another indication for a widespread acceptance of the increasing awareness around Serious Games opportunities. However, it complicates attempts to look at it from a scientific perspective towards identifying needs, requirements, gaps and research potential in a multi-disciplinary environment.

While there have been efforts to take a structured look at the various application areas for Serious Games (Table 1), identifying what differentiates a Serious Game from a game that is providing mere entertainment value seems daunting. What makes matters worse is that the aforementioned successes of recent activity games led to the adoption of the term Serious Games by marketers, attributing it to any possible positive impact the game play might have whether intended or not.

The following section is, therefore, an attempt to separate Serious Games for other types of games as well as to delineate Serious Games design considerations from coincidental similar design outcomes.

Drivers and resulting design features

In general, there are two main purposes for Serious Games: skill augmentation or behavioral change (e.g., Lieberman 1997). Both are equally valid, however create distinct requirements and approaches when it comes to design considerations and design features, especially due to the relative immaturity of the field.

Drivers for Serious Games targeting **skills augmentation** are the need for a more skilled workforce as well as safety improvements, thus addressing the weaknesses and shortcomings of a contemporary educational or training system and trying to exploit the strength of games for certain types of education and training. Consequently, design features include the manipulation of time, frequent repetition, the addressing of specific skills, and the gathering of related data towards proven mastery. Often the exposure to unacceptable risks or danger is an important design feature, introducing the stress, pressure and frustration required to make skills transfer effective.

Towards **behavioral change**, however, the drivers are the identification of needs *not* yet met by standard procedures or too costly to address by standard means. Consequently, other main drivers are peer-reviewed research as well as innovative and visionary individuals in academia, industry, government and non-governmental organizations, who are willing to take the risk to enter (and advocate) unchartered territory. While here too, the manipulation of time, frequent repetition and gathering of data are core design features, it is the focus on improving and maximizing self-efficacy that fundamentally changes the experience and game play.

Educational Games

Despite the previous postulation of Serious Games as effective media for education and training, there is a fundamental difference between Serious Games that educate by experience and Educational Games that communicate educational content in a game-like manner. Serious Games in this context develop connections between contributing artifacts, raise awareness, and are immersive in nature, whereas Educational Games develop specific skills, address and test specific knowledge, and are repetitive in nature. Not surprisingly, the latter are mostly developed for the public sector spear-headed by companies like Leap Frog and Vtech, while Serious Games for education are advocated by public institutions, business colleges and other non-profit organizations that are striving for a compelling experience with potential educational value.

Virtual Worlds

The emergence of massively multi-user online worlds or MMOW has led to the confusion of those virtual worlds with massively multiplayer online games or MMOG, which have in recent years successfully provided compelling and immersive game experiences to its audiences. However, MMOW such a Linden Lab's Second Life, Microsoft's Club Penguin or Numedeon's Whyville are platforms which certainly can be the interface to multiplayer distributed online Serious Games. With the freedom they are providing the user to interact, choose alternatives and provide input as well as their high level of instrumentation to give evaluation feedback to users and designers alike, they might actually be a reasonable basis for the earlier mentioned spectrum of life-related game activities that casual gamers are susceptible to. However, they shouldn't be confused with the game-experience focused, highly scripted and thus from a Serious Games perspective somewhat limiting MMOG such a Blizzard Entertainment's World of Warcraft, which trade participatory narrative for better (mostly visual) experience. Whether such distinction will hold up in the future needs to be seen since increasing online provision of game play and consumer input to game design and game scenarios might eventually overcome these disparities.

SERIOUS GAMES IN THE CONTEXT OF PERSUASIVE TECHNOLOGY

True Serious Games aim at having a *measurable* impact on the player without the player necessarily being aware of the intended impact. In fact, if the intended impact has not been established by the playing individual but by a higher authority such as a parent, physician, supervisor or educational institution, then there is a natural resistance or lack of buy-in, which will ultimately affect the game appreciation and thus experience. In recent years, the research field of persuasive technologies has emerged addressing particularly this challenge. How can technologies be designed that effectively change people's thinking and behavior without explicitly requiring the user working towards such change? B. J. Fogg's research (Fogg 2003) is studying technology persuasion along a functional triad as depicted in Figure 1.

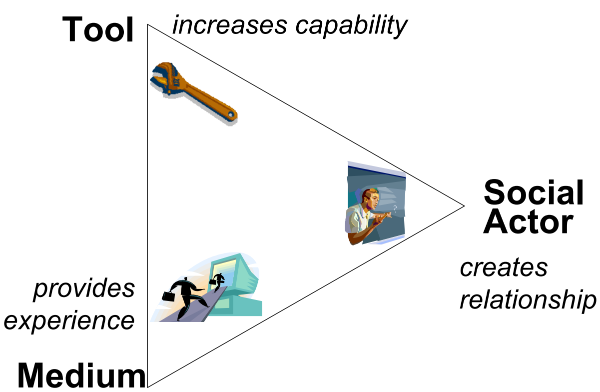


Figure 1: *Functional triad of persuasive technology (Fogg 2003).*

Based on this research, technology can be persuasive by any combination of increasing capability, creating a relationship or providing an experience. For instance, a *tool* can be persuasive by

- making target behavior easier to do;
- leading people through a process;
- performing calculations or measurements that motivate;

a *medium* can be persuasive by

- allowing people to explore cause and effect relationships;
- providing people with vicarious experiences that motivate;
- helping people rehearse a behavior;

and a *social actor* can be persuasive by

- rewarding people with positive feedback;
- modeling/demonstrating a target behavior or attitude;
- providing social support.

Applying this 'lens' of persuasive technology, one can easily see the strength of Serious Games in covering all three bases. Currently they are, however, predominantly depicted through implementations which provide experiences in which individual players are represented by social actors in the form of avatars and interact with other social actors in the of non-player characters (NPC). A more thorough discussion of video games in the context of persuasion has been published by (Bogost 2007).

Applying such methodology allows the evaluation of Serious Games in a way less tedious and restricted than trying to determine long-term behavioral change by means of short-term studies in controlled environments. While certainly less precise than through biometrical and physiological monitoring, such evaluations can be conducted before and throughout the design of the game rather than after its implementation.

OPPORTUNITIES AND CHALLENGES

Business processes in the traditional gaming industry follow business-to-consumer (B2C) models in the retail market or direct-to-consumer (D2C) models through emerging online distribution channels. In contrast, the significantly higher fragmentation of the Serious Games market requires the availability of very specialized distribution channels thus calling for business-to-business (B2B) models. Moreover, the traditional gaming industry's success is based on the common public recognition that games are entertaining. Therefore, development, marketing and distribution only need to cater to the preferences of various demographics to target significant sales success with large audiences in the tenth and hundreds of millions.

Serious Games, on the other hand, lack both, the public acceptance of being generally a value-add as well as large audiences to design and market towards. The proof of value-add has so far been stifled by the need for scientific evidence that a Serious Game had a skill augmentation or behavior change impact on its players. Furthermore, there were less people with a particular self-motivated need for improvement or change than with the inherent desire to have fun. This required the availability of dedicated sales channels and organizations and therefore a B2B model. Such model lacked so far though the market size, on the one hand, and a robust and diversified value chain on the other to meet all audiences' needs and thus exploit all commercial opportunities.

Times might, however, be changing in favor of Serious Games. The financial, educational, and healthcare crises worldwide are affecting large audiences and are creating the public awareness that change is inevitable and must be a combination of public and individual effort and responsibility. Associated shortfalls in schools, the workplace, social communities and individual behavior such as poor education, mass layoffs, bankruptcy, physical inactivity, malnutrition, and drug-abuse are affecting large demographics. Ironically, those problems affecting the masses are taking the stigma of resulting individual problems such as depression, obesity, illiteracy, unemployment, and poverty. Simultaneously, the field of Serious Games research and development has matured and was able to produce some compelling highly publicized evidence on particular game interventions, which created public and individual awareness around the potential of Serious Games (e.g., Baranovski et al. 2008, Kato et al. 2008). Consequently, several private and non-government organizations have invested in rapidly evolving the field of Serious Games from a mere academic field of study to a profitable industry. In anticipation, the established game industry has realized the trend and is beginning to start strategic directions as previously discussed in order not to be left out.

A remaining challenge for Serious Game development is related to the fact that purposeful play is something very personal and requires careful tailoring to **individual needs, preferences and contexts**. For it to have a measurable impact that translates well to real-world proficiency and behavioral change requires the experience to be very realistic. Game development over the years got away with fictional representations of a physical environment, approximations of physical behavior and crude as well as imprecise input devices. In contrast, many Serious Games applications might require a higher degree of visual, tactile, olfactory or other-sensory realism than commodity gaming hardware and software had to provide. Serious Games also might have to tie to existing real-world equipment in order to facilitate higher and quicker knowledge and experience transfer. At the same time, Serious Games have to still be able to compete with the appeal and experience of entertainment games which don't have those additional constraints.

Finally, Serious Games have the requirement to capture and **measure** the various facets of individual game play that document a player's progress towards the embedded purpose of the game. Availability of such data is not only crucial to integrate Serious Games seamlessly with other programs aimed at improving skills or self-efficacy. It is also a prerequisite towards improving the game itself and further document the business viability of Serious Games to various stakeholders. However, the fact that true Serious Games as defined above might not explicitly address these purposes but rather persuade players through the associated experience, functionality, or communication can make such data hard to capture. Such data is not directly tied to game design or game play but rather needs to be inferred from the players reaction.

All these factors increase the **cost** for developing Serious Games, on the one hand, and are creating the risk for Serious Games to remain an academic field of study struggling to prove ultimate viability and consequently handicapped by lengthy, tedious and very narrowly focused scientific research endeavors resulting in mere proofs of concept.

THE SCIENCE PERSPECTIVE

The following discussion is trying to identify fields of study which – based on the previous considerations of Serious Games opportunities, challenges, and role in the context of persuasive technology – seem crucial in order to advance the science in support of increasing market momentum and discovering new application areas. The following list is certainly not exclusive and advancement in science and research in many other areas will be necessary to reach for the vision of life-accompanying Serious Game play for the benefit of individuals and their societal and environmental context. At the same time, this list is trying to avoid areas that are already heavily covered by traditional game development research. Progress in those areas is to be expected independent from the particular requirements for Serious Games.

Computer Graphics & Computer Vision

The need for increased realism and authenticity at reasonable cost requires new approaches that are able to rapidly capture and display existing environments and artifacts and integrate them with the game play. The field of Computational Photography (e.g., Tumblin and Raskar 2006) has recently matured to a discipline that promises to provide inclusive photorealistic experiences from real-world imagery while Google's Streetviewer has already begun to provide a glimpse at its commercial potential. At the same time camera technology has progressed to a point that the 3D reconstruction from cameras will soon become a commodity. Sony's eyeToy and Microsoft's Natal were just the beginning of exploring 3D camera recognition for interaction purposes but it can be expected that this progress will continue towards other areas of camera recognition applications. However, much research is needed on how such technology can be employed in mobile and outdoor environments, and can drive or integrate with games that run even on mobile and low-cost gaming devices.

Human-Computer Interaction

In the days of Virtual Reality (VR) research much emphasis was put on optimizing the realism by means of multimodal input and multisensory feedback (e.g., Steuer 1992). The cost and immaturity of technology then as well as lacking market demand for VR applications hampered the progress in these areas. However, today's market penetration of commodity gaming platforms which exhibit much higher graphics and processing performance than their expensive special purpose predecessors makes scientific research in **multimodal input** and multisensory feedback again an academically worthy endeavor with significant commercialization potential. Controller-free input such as demonstrated by Sony's eyeToy, Microsoft's Natal, gesture recognition such as underlying the Wii input technology, and face and voice recognition such as also promised in Natal need to be combined with outcomes of the research fields of Affective Computing and Intelligent Interaction (e.g., Paiva et al. 2007) capturing emotions, attitudes and unconscious behaviors by means of computer vision as well as different biometric and physiological sensors.

Multisensory feedback addresses the need of the player to receive realistic system response that corresponds with the game application as well as the underlying game purpose of a Serious Game in order to have the potential to invoke success in skills augmentation or behavioral change. While aural feedback has reached especially through the introduction of spatial audio an acceptable level of maturity, other senses have continued to suffer from neglect. Tactile feedback so far is only available as rudimentary vibration of game controllers, haptic devices are only available for industrial or highly specialized applications (e.g., Bayonat et al. 2006), olfactory displays are still very much limited to academic Virtual Reality applications (e.g., Yanagida 2004), and so are other types of sensory renderings such as heat, wind, and humidity. Revisiting the research in those technologies in the context of Serious Games applications as well as developing a methodology on how to synchronize the various modalities appropriately in the context of game play and interaction seems to be an equally necessary and exciting challenge; not to mention the challenge of taking those modalities then to the low-cost and mobile applications space.

Modeling, Simulation, and Artificial Intelligence

In order to have an impact, Serious Games must be more concerned than traditional games with creating an accurate model of the player. This is in order to better tailor the game experience to the player's needs and preferences, including potential NPCs to accurately communicate with and hopefully persuade the player. Such player-adaptive experiences require a level of artificial intelligence to appropriately respond, anticipate and direct the human user's actions that is not yet available in current systems. If, however, an additional goal of modeling the players is to reflect themselves in the game (such as in first-person genres and RPG's) then we additionally need much more accurate simulations of human's in the game. This statement does not only refer to the visual appearance, physical behavior or physiological composition of humans as being the focus of much computer graphics and medical simulations research these days. It moreover includes the aspect of personalities, social behaviors, emotions, attitudes and complex interactions of body, mind and spirit (cf. e.g., van Lent and Swartout 2007) which – in nature – have in return an influence on the visual and physiological representations of individuals. Only if such complex models of overall human existence can be established, can we target effectively aforementioned societal problems through Serious Games such as health including obesity and depression.

Ubiquitous Narratives and Transmedia Game Play

The ubiquity of gaming devices is addressing the increasing request by players to game at any time and place of their choice. From home-bound game consoles and PCs to handheld devices and mobile phones, gaming has become an integral component of everyday life. Consequently, researchers have started to investigate game play and storytelling that transcend game platforms and are, therefore, able to more continuously engage players in the narrative thus providing a more immersive and compelling experience. Under the term transmedia storytelling (e.g., Jenkins 2006), experiences are being designed that are not only replicated across various media but continue across different media drawing from the strength of each individual media channel. Alternate Reality Games (ARG) specifically addresses the aspect of gaming in the real world supported by narratives involving massive amounts of players across various different media (e.g., McGonigal 2005, Kim et al. 2008). Supported by complementary mobile applications aimed at integrating synthetic game content and real-world experiences such as Mobile Augmented Reality Games (e.g., Broll et al. 2008) and other mobile graphics technologies (e.g., Mobile Graphics 2008), these research directions hold promise with respect to two important challenges for Serious Game development: On the one hand, they promise to provide the foundation for business models as basis for Serious Game play by tying it to real-world media and business infrastructures (McGonigal 2008). On the other hand, they provide the growing movement of social gaming a platform to support infinite number of social groups and their objectives. Lastly, they provide to the increasing number of casual gamers involved in social gaming a means to address and engage around their personal passions – not through continuous repetition but rather through day-long continuous experiences.

Evaluation & Assessment

The increased demand for measurable impact in Serious Games in combination with an infinite number of applications and associated individualized goals and purposes make traditional effectiveness studies highly impractical beyond pathbreaking proofs of concept. In addition, the increasing ubiquity of game play makes it highly impracticable to conduct field studies for all possible game play contexts. Possibly the integration of emerging small, inexpensive and even disposable biometric sensors could provide the means for remote monitoring in context as well as in-situ adaptation of game play to player performance and behavior. In addition, advanced sensor technology to capture the corresponding context and circumstances as well as the networking and communication infrastructures to support such adaptive ubiquitous Serious Game play provide further challenges for science and engineering worth addressing.

Standardization

Serious Games more than traditional games have the need for a high degree of standardization in order to gain the ability for personalized game play. Based on the previously discussed models of human personality and behavior, mutually exchangeable components of narrative, game play, interface, and platform need to be adapted towards optimizing the Serious Game's underlying purpose to the player. At the same time, the cost of compelling game development in general and specifically Serious Games as addressed before requires the establishment of interfaces and interoperability standards in order to maximize reuse. Ideally any commercial entertainment game could be used as the basis for a suite of Serious Games by just plugging in the additionally needed components for audience-tailored skill development or behavioral change. Such capability would create a win-win situation between traditional game publishers (who would be able to increase sales) and Serious Games developers who could significantly reduce cost. In a similar way as Computer Graphics took off after standardization efforts like PHIGS, GKS and CORE led to ultimately OpenGL, standardization successes with respect to Artificial Intelligence, Human Modeling, game mechanics (for different game genres) and narrative design would likely provide a significant accelerator to the field of Serious Games.

Education and Collaboration

Many of the scientific and research challenges outlined above can only be successfully addressed through two fundamental processes.

Interdisciplinary research and education will ensure that all the various disciplines required will work together to accomplish the common goal of effective Serious Games development, including but not limited to Computer Science, Human Factors, Psychology, Human-Computer Interaction, Arts, Anthropology, Education, and Sociology. It will not be enough to establish multi-disciplinary collaborations and curricula as widely pursued (see e.g. Zyda 2007). Rather than the various disciplines focusing on their particular aspect and perspective of the common goal it is required to create truly transdisciplinary curricula and research experiences that ensure knowledge generation and transfer.

Likewise, Serious Games are much dependent on the particular application area which requires practitioners and industry to join in and provide the domain expertise to ensure accelerated innovation transfer from academic research to market application.

CONCLUSIONS

The emergence of social gaming and the corresponding increasing number of casual gamers entering the market for gaming purposes beyond mere entertainment is providing new opportunities to Serious Game development. Especially by addressing global and societal relevant problems like education, health, employment and politics, the field of Serious Games will not only be able to overcome its cost and market size challenges but provide attractive market potential for the traditional game industry which is at the mercy of fluctuations in the economic wellbeing of its player constituents. Along the way are a variety of scientific and research challenges that need be addressed, yet by themselves can already contribute to solving those societal problems through economic and educational stimulation. One prerequisite is, however, a new type of transdisciplinary partnership between academic disciplines and application industries leading to a whole new intellectual and practitioner workforce as well as market-stimulating products and services in Serious Games.

REFERENCES

Alhadeff, E. 2007. "Serious Games – A sizeable market". In Blog *Future-making Serious Games.* URL: http://elianealhadeff.blogspot.com/ (last visited 06/26/09).

Baranowski, T., R. Buday, D. I. Thompson, and J. Baranowski 2008. "Playing for real: Video games and stories for health-related behavior change." *American Journal of Preventative Medicine.* Jan; 34(1): 74-82.

Bayonat, S., M. Garcia, C. Mendoza, and J. Fernindez 2006. "Shoulder Arthroscopy Training System with Force Feedback." In MEDVIS Proc. of the Intl. Conf. on Medical Information Visualization, Biomedical Visualization. 71-76.

Broll, W., I. Lindt, I. Herbst, J. Ohlenburg, A. Braunand R. Wetzel 2008. "Towards Next-Gen Mobile AR Games," *Computer Graphics and Applications*, Vol. 28, No. 4, July/August 2008, IEEE Computer Socciety, 40-48.

Bogost, I. 2007 *Persuasive Games: the Expressive Power of Videogames.* The MIT Press.

Fogg, B.J. 2003. *Persuasive Technology: Using computers to change what we think and do.* Morgan Kaufmann Publishers, Amsterdam, Boston.

Jenkins, H. 2006. *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.

Kato, P. M., S. W. Cole, A. S. Bradlyn and B. H. Pollock, 2008. "A Video Game Improves Behavioral Outcomes in Adolescents and Young Adults With Cancer: A Randomized Trial." In *Pediatrics*, Vol. 122 No. 2 August 2008, e305-e317.

Kim, J. Y., J. P. Allen, and E. Lee 2008. "Alternate Reality Gaming," *Communications of the ACM*, Vol. 51, No. 2, February 2008, ACM, 36-42.

Lieberman, D. 1997. "Interactive Video Games for Health Promotion: Effects on Knowledge, Self-Efficacy, Social Support and Health." In *Health promotion and interactive technology*, Richard L. Street, William R. Gold, and Timothy R. Manning, eds. Lawrence Erlbaum Associates, Mahwah, NJ. 103-119.

McGonigal, J. 2005. "Supergaming! Ubiquitous Play and Performance for Massively Scaled Community." *Modern Drama.* Special Issue on Technology. Ed. W.B. Worthen. 48:3, 471-491.

McGonigal, J. 2008. "Making Alternate Reality the New Business Reality." Op-Ed. *Harvard Business Review.* Special Issue: Top 20 Breakthrough Ideas for 2008. February 2008.

Mobile Graphics 2008. *Computer Graphics & Applications*, Special Issue on Mobile Graphics, Vol. 28, No. 4, July/August 2008, IEEE Computer Sociaty.

NPD Group. 2009. Entertainment Connection. URL: http://www.npd.com/lps/entertainment_connection/ (last visited 06/26/09).

Paiva, A., R. Prada and R. W. Picard (Eds.) 2007 *Affective Computing and Intelligent Interaction*. Proc. of the Second International Conference, ACII 2007, Lisbon, Portugal, September 12-14, 2007. Lecture Notes in Computer Science, Image Processing, Computer Vision, Pattern Recognition, and Graphics , Vol. 4738, 2007, XVIII, ISBN: 978-3-540-74888-5

Sawyer, B. and P. Smith. 2008. "Serious Games taxonomy" *Presentation at the Serious Games Summit at the Game Developers Conference*, San Francisco, CA. URL: http://www.seriousgames.org/presentations/serious-games-taxonomy-2008_web.pdf (last visited 06/26/09).

Steuer, J. 1992. "Defining Virtual Reality: Dimensions Determining Telepresence." In *Journal of Communication*, Vol. 42, No. 2. Oxford University Press.

Tumblin, J. and R. Raskar 2006. "Computational Photography." State of The Art Report (STAR) in *ACM/Eurographics 2006*, Sept 1, Vienna, Austria. 25 pages.

van Lent, M and W. Swartout 2007, "Games: Once More with Feeling". *Computer*, IEEE Computer Society, Vol. 40, No. 8, August 2007, 98-100.

Yanagida, Y., S. Kawato, H. Noma, A. Tomono, and N. Tesutani 2004. "Projection based olfactory display with nose tracking." In Proc. *IEEE Virtual Reality 2004*, 27-31 March, 2005, 43-50.

Zyda, M. (ed.) 2007. "Creating a Science of Games". *Communications of the ACM*, Special Issue on Creating a Science of Games, ACM, New York, 26 pp.

BIOGRAPHY

Dr. L. Miguel Encarnação has spent fifteen years in academic and applied research in the areas of human-media interfaces, virtual and augmented reality, digital storytelling, advanced distributed learning and Serious Games. In the area of Serious Games R&D his research teams have over the years worked with several academic, industry, government and non-government organizations including the US DoD, UN preparatory commission, Hasbro, SAP, Mystic Aquarium, University of Southern California and Georgia Institute of Technology.

As the Director for Emerging Technology Applications at Humana's Innovation Center his particular application areas of interest currently focus on health and wellness including entertainment, education, and social engagement.

Encarnação received his PhD in Computer Science from University of Tübingen, Germany and has been serving as adjunct professor of Computer Science at the University of Rhode Island since 1998.

Dr. Encarnação is the author or co-author of numerous contributions to peer-reviewed journals and conferences, has co-chaired symposia and conferences on computer graphics technologies and applications, and has made contributions to books on computer graphics education and programming.

He is the Associate Editor in Chief of IEEE Computer Graphics & Applications, a member of the editorial boards of the International Journal of Technology & Human Interaction and the International Journal of Virtual Reality, and a frequent reviewer for various technical journals as well as the U.S. National Science Foundation. He is a senior member of the ACM, an ACM SIGGRAPH pioneer, as well as a member of ACM SIGGRAPH, ACM SIGCHI, the IEEE Computer Society, Eurographics, and the German Society of Computer Science (GI).

PARAMETERIZED GENERATION OF AVATAR FACE DATASET

Justin N Oursler, Mathew Price, and Roman V. Yampolskiy
Computer Science and Engineering
JB Speed Hall, 106
University of Louisville
Louisville, Kentucky, USA
jnours01@louisville.edu, mahatah@gmail.com, roman.yampolskiy@louisville.edu

KEYWORDS
Artimetrics, Avatar, Dataset, Face, Second Life.

ABSTRACT

Virtual communities such as *Second Life* and *Entropia Universe* are quickly becoming the next frontier of cybercrime. With GDPs of virtual economies of persistent worlds approaching billions of dollars it is necessary to develop tools for protection of virtual environments similar to those utilized to secure real world's infrastructure, such as biometrics security systems. The first step in the development of such security methodologies is wide availability of datasets necessary for training and testing of those biologically inspired systems. In this paper we present a manual and automated approach to generation of parameterized datasets containing facial images of avatars representing typical entities in the virtual worlds. Our work on making such standardized datasets makes it possible to develop novel security systems which function just like biometric recognition systems, but can be applied to recognition and verification of non-biological entities.

INTRODUCTION

Virtual economies of persistent worlds have Gross Domestic Product (GDP)s approaching billions of dollars, far exceeding GDPs of some real countries. Virtual crime results in real financial damages which will only worsen if measures are not taken to secure authentication and communication in virtual reality worlds. In the virtual world increasingly populated by non-biological characters there are just no existing techniques for identity verification of intelligent entities other then self-identification.

Artimetrics, which is defined as the science of recognition and verification of intelligent software agents, domestic and industrial robots and other non-biological entities aims to address this problem. This future oriented sub-field of security has broad applications (Yampolskiy, 2008). Artificially Intelligent programs are quickly becoming a part of our everyday life. Virtual assistants, shopping bots, and smart search engines, to give just some examples, are used daily by millions of people. As such programs become closer in their abilities and intelligence to human beings, the need will arise to recognize and verify the identity of such entities just like it is necessary to authenticate the identity of people. With respect to artificial agents, such reasons include but are not limited to (Yampolskiy, 2008):

- Preventing malicious intelligent software from obtaining access to information or system resources and granting it to authorized agents and by doing so improving security of virtual communities, social networks, and country's cyber-infrastructure especially vulnerable in the post 9/11 world.
- Finding out which agent has performed a given task in case a number of possible alternatives exist, either for demanding responsibility or assigning reward.
- Securing interaction between different pieces of intelligent software or between a human being and an instance of intelligent software/robot.
- Determining who has the authorship rights to the results of computation and creative output produced by an AI entity.
- Identifying semi-autonomous software tools used by hackers to attack systems and networks.
- Making it possible for scientists in fields as diverse as biology, communications, and e-business to securely work with intelligent assistants, and robots.

PREVIOUS WORK

Artificially generated datasets are a common approach in development of biometric security systems. Once a system for producing simulated data is developed it is fast and cheap to produce large quantities of high quality biometric data without any privacy or security concerns to worry about

(Yampolskiy, 2008). Such data can be used for testing newly developed biometric systems, benchmarking well developed security systems, testing scalability of authentication systems or for certification of commercially available packages. Production of synthetic biometrics allows researchers to better understand individuality of biometric patterns and allows parametric sensitivities of algorithms to be investigated in greater detail (Orlans, 2004; Ma, 2005; Makthal, 2005).

Biometric data comes in four different formats: images, feature measurements, matching scores and decision data. Each type can be potentially simulated for experimental purposes (Ma, 2005).

- **Images** This is the best researched type of synthetic biometric represented by the raw image as it comes out from a digital scanner or a camera and depicts a two-dimensional view of the body part/structure in question (Ma, et al. 2005). Fingerprint, face and iris images are examples of the most frequently produced artificial data (Cappelli, 2000; Cui, 2004; Ayers, 2004; Cappelli, 2002; Sumi, 2006; Makthal, 2005).
- **Matching Scores** Similarity functions attempt to measure closeness of the input feature vector to the one stored in a template. A decision on user verification or recognition is produced by the system based on such a measurement. Generation of matching scores may be valuable for testing a model for the actual distribution of matching scores for a particular biometric (Wein, 2005; Ma, 2005).
- **Decision Data** This is the highest level of biometric data which can be generated. All biometric systems produce a binary decision in terms of authenticating the user or not. Synthetic decision data can be used to estimate error rates produced by a biometric system (Schuckers, 2004).
- **Feature Measurements** Extraction of statistically valuable information from the raw images produces feature measurements such as the number of minutiae points in a fingerprint or the inter-eye distance in face recognition.

Currently well researched synthetic biometric data generation approaches exist for many physical biometrics such as fingerprint (Cappelli, 2000; Ayers, 2004; Cappelli, 2002), face (Sumi, 2006), and iris (Cui, 2004; Makthal, 2005). Additionally, ability to generate many others comes from other fields such as computer graphics for face (DeCarlo, 2008; Gao, 2000) and iris (Lefohn, 2003), speech synthesis for speech (Taylor, 1999), document image analysis for signatures (Oliveira, 1997), and speech processing for lips (Reveret, 1998).

Virtual World Selection

For the purposes of avatar facial data generation, various virtual worlds and avatar creation software were considered based on the needs of this project. These included:

1) Mutable attributes to avatar facial features
2) Ability to view avatar from multiple angles
3) Selection of contrasting facial features in generating new avatars
4) General ease of use and versatility

While many virtual worlds such as *Entropia Universe* and avatar creation software such as *Poser* were considered, the Massively Multiplayer Online Role Playing Game (MMORPG) *Second Life* (Second Life, 2009) was found to best fit the above criteria. Second Life, a three-dimensional, virtual world where the Internet community creates avatars to interact with one another in real-time, provided the largest, quickest, easiest to use, and most versatile set of avatars and creation materials of all software considered.

Second Life's benefits were many in consideration for the creation of a dataset, but four attributes stood out among all others. First, dozens of physical facial attributes, such as length and right-to-left symmetry, could be adjusted to create a truly unique avatar. Secondly, Second Life implemented an avatar randomizer which can create a new, unique avatar with the press of a button. Thirdly, camera pan, tilt, and zoom are controllable by the user which makes it possible to gather several angles from the same avatar. Finally, Second Life allows the user to manipulate environmental elements through the use of an in-game scripting language, adding versatility to the collection of data.

DEVELOPED METHODOLOGY

Manual Approach Each unique avatar face was randomly generated and saved via the Gadwin PrintScreen (Screen, 2009) application which allowed us to quickly capture the screen and save it in the desired directory as a certain file type. The following steps were taken to capture the profile of each avatar:

1) Open the Appearance Menu and randomized each category of the avatar's physical appearance, ending with the eyes. By ending

with the eyes, the focus would automatically go to the avatar's face, allowing consistency in the dataset.

2) Adjust the view of the screen, putting the avatar in the center of the screen, capture the screen, and save the image in the dataset directory with the appropriate filename.
3) Rotate the camera to the left where the avatar is still facing forward but we can still capture a left-handed angle in between this view and the center view. Then, capture as before.
4) Rotate the camera in between the far left view and the center and then capture.
5) Rotate the camera to the far right, again, so the avatar is still facing forward but where a capture can still be performed between this view and the center view. Then, capture the screen.
6) Rotate the camera between the center view and far right view and capture the screen.
7) Focus the camera on the center again, but then rotate the camera to a view below the avatar, with the avatar still facing forward. Then, capture as before.
8) Rotate the camera straight through the center to a view above the avatar with the avatar still facing forward. Then, capture the screen.

These eight steps generated one image set for each avatar. Problems with this approach mainly consist of possible human error and irregularities in the collected data. Though a possible problem, this could also be viewed as a benefit since a true application of our research will likely work with noisy data. Giving the future algorithms for avatar face detection and recognition a random selection of views makes them better equipped for real-world applications. Also, manually producing avatar facial database is a very time consuming task, and so is not a feasible approach for larger size datasets.

The dataset generated by the manual approach consists of one hundred different avatars with seven pictures from different angles, totaling seven hundred images. The images were captured as Tagged Image File Format (TIFF) (TIFF, 2009) images at a resolution of 1280 X 1240, resulting in each image being 3,843 KB in size (see Figure 1-4). The TIFF format was used because it is a flexible format that allows for customization in its tags (image information) and it is uncompressed, maintaining the quality of the image. The first fifty avatars (three-hundred fifty images) consist of randomly generated male avatars and the last half of the images is of female avatars. An angle from the a) front, b) far left, c) mid left, d) far right, e) mid right, f) bottom, and g) top, is captured for each avatar. The images are named in a consistent format; stating the program, character, and angle. For example, the image SL-051b.tif, is a capture using *Second Life* (SL) of the 51st character (051) from the far left (b) that happens to be female because it is in the last half of the dataset.

Automated Approach For a second approach, we designed and implemented a scripting technique to automate the above process. Using the programming language AutoIt (AutoIt, 2009) as well as a scripting language native to *Second Life*, better known as Linden Scripting Language (LSL) (LSL, 2009), a successful generation of random avatars was achieved. The following is a walkthrough of this process for the creation of one hundred randomly generated avatars:

1) Using the scripting language AutoIt, it was possible to simulate key presses and mouse control in a Windows environment. During the first run of the AutoIt script, simulated keyboard commands are used to circle the *Second Life* camera around the avatar such that the front of the avatar's face is exposed.

2) The script is paused and requests the user to center the avatar's face with the horizon using the movement control. This is only needed on the first run and constitutes the last interaction with the user.

3) The AutoIt script then activates the LSL script by clicking on a button attached to the avatar's hub.

4) The LSL script locks the *Second Life* camera's position and rotation as well as controls from the game's automated functions (such as camera changing on clicking).

5) The AutoIt script then takes a screen shot of the avatar using the *Second Life* tool "screen shot". The script then labels avatar "Avatar 'x' face 'y'", where x corresponds to the number of avatar created (1 - N) and y corresponds to the screen shot for that avatar (1 – 10).

6) The script then zooms into the avatar's face before taking another screen shot and using the same labeling system as in step 5.

7) The AutoIt script then rotates the camera at eight specific angles (upper left, center left, lower left, upper center, lower center, upper right, center right, and lower right) taking screen shots at each.

8) The script then selects "edit", then "appearance", bringing up the avatar editing tool. From here the script randomizes a body for a new avatar. Body

height, torso length, and leg height all must be set to 50% in order to preserve the camera angle, which is done automatically by the script.

9) The AutoIt script then clicks on the body parts sub menu items “skin”, “hair”, and “eyes” randomizing each of them as they are entered. The save “all button” is pressed, saving the avatar to begin the screen shot process again.

10) The script zooms away from the avatar before taking the new avatar's center body screen shot.

After step 10, the AutoIt script restarts at step 7 until one hundred avatars are created and one thousand images have been taken. A sample segment of Autoit source code responsible for GUI interaction is given below:

```
Func snapshot ($picture)    dim $picture
mouseClick("Left", 440, 756, 1)   ;snapshot button
sleep(2000)
mouseClick("Left", 102, 296, 1)   ;save button
sleep(3000)
send("{DOWN}{ENTER}")
findname($picture)                    EndFunc
```

The datasets generated by the scripted approach consists of avatars with ten pictures from different angles. The images captured are in the Portable Network Graphics (PNG) (PNG, 2009) format at a resolution of 1024 X 768 resulting in each image being between 110KB and 450KB in size. One upper body picture is taken as well as nine facial pictures, all differing in angles. These angles include the top, center, and bottom of the left, center and right side of each avatar's face. The images are named in a consistent format; stating the program, gender, avatar number, and angle. For example, the image “SecondLife Male Avatar 4 gesture 5.png” refers to the image of an avatar that looks like a male character, the fourth in the dataset, and the fifth picture taken in this avatar's set of 10 (see Figures 5-8). The gender of the avatar is dependent upon the user’s selection at the beginning of the process.

CONCLUSIONS AND FUTURE RESEARCH

Now that we have a completed dataset of avatar faces, our next step is to utilize the dataset in different security related experiments. We will first experiment with avatar face detection and then face recognition. Face detection is the process of determining whether or not there is a face in the image and precisely locating it. The experiments involving face detection will consist of utilizing and improving existing face detection algorithms currently used for human face detection in biometric applications. After we are successfully able to detect faces we will begin experimentation with face recognition algorithms. Face recognition aims to identify if a specific face is included in the dataset of previously enrolled faces. Again, experiments will consist of applying existing biometric algorithms and improving their design until we can successfully authenticate avatars via face recognition. First, our algorithms will only be tested with still photos taken from in-game screenshots. This allows us to keep the problem simple for testing purposes. Our ultimate goal is to be able to utilize an algorithm that can detect and recognize an avatar in a three-dimensional environment like in the virtual world *Second Life*.

Being able to identify software agents not just by certain specific codes (ID numbers, serial keys, etc.), but also visually adds extra security to the authentication process. If a person, through an avatar, is wanting to exchange money for goods and services in an application such as *Second Life*, the program can authenticate that avatar by not only their factual information but also by the way they are expected to look, which provides a second level of authentication or alternatively allows for decentralized (serverless) virtual government with vibrant economy. The security in our real world is ever changing and the security in our virtual world must do the same in order to keep up with the real crimes committed via wires and waves. No longer do we only need to worry about human faces and identities but non-human entities with real abilities to cause harm.

BIOGRAPHIES

Justin N. Oursler is currently studying Computer Engineering and Computer Science at the J.B. Speed School of Engineering at the University of Louisville. Justin is an active member of the Sigma Chi fraternity and representative of the Speed School Student Council.

Matthew E. Price holds an undergraduate degree in Justice Administration from the University of Louisville. He is currently pursuing an MS degree in Computer Science and Engineering from the University of Louisville. Mr. Price's interests include computer forensics and biometric systems.

Roman V. Yampolskiy holds a PhD degree from the department of computer science and engineering at the University at Buffalo. He is currently an assistant professor at the University of Louisville. Dr. Yampolskiy’s main areas of interest are computer security and biometrics. Dr. Yampolskiy is an author of over 40 publications including multiple books.

Fig. 1. Manually collected image SL-002a.

Fig. 2 Manually collected image SL-024c.

Fig. 3. Manually collected image SL-054e.

Fig. 4. Manually collected image SL-095g.

Fig. 5. Automatically collected image Second Life Male Avatar11 gesture3_001.

Fig. 6. Automatically collected image Second Life Male Avatar11 gesture5_001.

Fig. 7. Automatically collected image Second Life Male Avatar22 gesture2_001.

Fig. 8. Automatically collected image Second Life Male Avatar22 look7_001.

REFERENCES

Ayers, T., R. Federkeil, et al. (2004). Modeling Fingerprints: Components for the Task of Synthesis. Proc. BT2004 Int'l Workshop on Biometric Technologies, AB, Canada.

Cappelli, R., D. Maio, et al. (August 2002). Synthetic Fingerprint-Database Generation. in proceedings 16th International Conference on Pattern Recognition (ICPR2002), Quebec City, Canada.

Cappelli, R., D. Maio, et al (2000). Synthetic Fingerprint-Image Generation. 15th International Conference on Pattern Recognition (ICPR'00).

Cui, J., Y. Wang, et al. (23-26 August 2004). An Iris Image Synthesis Method based on PCA and Super-resolution. IEEE International Conference on Pattern Recognition, United Kingdom.

DeCarlo, D., D. Metaxas, et al. (1998) An Anthropometric Face Model using Variational Techniques. SIGGRAPH '98.

Gao, W., J. Yan, et al. (September 2000). Individual 3d face synthesis based on orthogonal photos and speech-driven facial animation. In IEEE International Conference on Image Processing (ICIP), Vancouver, Canada.

Jermyn, I., A. Mayer, et al. (August 23-36, 1999). The Design and Analysis of Graphical Passwords Proceedings of the 8th USENIX Security Symposium, Washington, D.C.

Lefohn, A., R. Caruso, et al. (November 2003). An Ocularist's Approach to Human Iris Synthesis. IEEE Computer Graphics and Applications.

Ma, Y., M. Schuckers, et al. (October 2005). Guidelines for Appropriate Use of Simulated Data for Bio-Authentication Research. 4th IEEE Workshop on Automatic Identification Advanced Technologies (AUTO ID), Buffalo, NY.

Makthal, S. and A. Ross (September 2005). Synthesis of Iris Images using Markov Random Fields. in Proc. of 13th European Signal Processing Conference (EUSIPCO), Antalya, Turkey.

Oliveira, C. d., C. A. A. Kaestner, et al. (November 1997). Generation of Signatures by Deformations BSDIA'97, Curitiba,Brazil.

Orlans, N. M., D. J. Buettner, et al. (2004). *A Survey of Synthetic Biometrics: Capabilities and Benefits*. International Conference on Artificial Intelligence (IC-AI'04), CSREA Press.

Reveret, L. and C. Benoit (Dec. 4-6, 1998). A New 3D Lip Model for Analysis and Synthesis of Lip Motion in Speech Production. Proc. of the Second ESCA Workshop on Audio-Visual Speech Processing, AVSP'98, Terrigal, Australia.

Schuckers, M. E., A. Hawley, et al. (2004). A comparison of statistical methods for evaluating matching performance of a biometric identification device- a preliminary report. Proceedings of SPIE Conference on Biometric Technology for Human Identification, Orlando, USA.

Sumi, K., C. Liu, et al. (Jan. 2006). Study on Synthetic Face Database for Performance Evaluation. International Conference on Biometric Authentication (ICBA2006).

Taylor, P. and A. Black (1999). Speech synthesis by phonological structure matching. In *Eurospeech99*, Budapest, Hungary.

Thorpe, J. and P. v. Oorschot (December 6-10, 2004). Towards Secure Design Choices For Implementing Graphical Passwords. 20th Annual Computer Security Applications Conference, Tucson, Arizona

Wein, L. M. and M. Baveja (May 2005). Using Fingerprint Image Quality to Improve the Identification Performance of the U.S. Visitor and Immigrant Status Indicator Technology. In Proceedings of the National Academy of Sciences.

Yampolskiy, R. V. and V. Govindaraju (2008). Behavioral Biometrics for Verification and Recognition of Malicious Software Agents. Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VII. Orlando, Florida.

Yampolskiy, R. V. and V. Govindaraju (2008). Generation of Artificial Biometric Data Enhanced with Spatial-Temporal and Environmental Information. Biometric Technology for Human Identification V. Orlando, Florida.

Second Life | What is Second Life? [cited 2009 March 01]; Available from: http://secondlife.com/whatis/.

Screen Capture and Print Screen software. [cited 2009 March 01]; Available from: http://www.gadwin.com/printscreen/.

AutoIt Script Home Page. [cited 2009 March 01]; Available from: www.autoitscript.com/autoit3/index.shtml.

LSL Wiki: HomePage. [cited 2009 March 01]; Available from: http://www.lslwiki.net/lslwiki/wakka.php?wakka=HomePage.

TIFF. [cited 2009 March 01]; Available from: http://partners.adobe.com/public/developer/tiff/index.html.

PNG (Portable Network Graphics) Home Page. [cited 2009 March 01]; Available from: http://www.libpng.org/pub/png/.

INTEGRATING FACIAL, GESTURE, AND POSTURE EMOTION EXPRESSION FOR A 3D VIRTUAL AGENT

Shawna C.G Tan and Alexander Nareyek
Department of Electrical and Computer Engineering
National University of Singapore
10 Kent Ridge Crescent,
Singapore 119260
E-mail: yellowoysters@gmail.com, elean@nus.edu.sg

KEYWORDS
Virtual characters, virtual agents, non-verbal expressions, emotion expressions.

ABSTRACT

Facial expressions, gestures, and body postures can portray emotions in a non-verbal way. These methods are frequently employed by actors in theatrical plays or in movies, or even by virtual characters such as those found in computer games, animated storybooks, and website e-assistants. Signals for emotion expressions (“cues”), such as a raised fist and narrowing of the eyes, substantially influence the viewers’ assumptions on the emotional state of the person portraying it. To enable humans to recognize emotions of virtual characters, the characters’ cues must be portrayed according to the human counterparts.

In this study, we first look at existing systems for synthesizing cues for facial expressions, gestures and body postures. This is followed by examining the emotion recognition problems that arise from utilizing the various systems. Lastly, the systems are integrated together and the implications that arise from the integration are analyzed.

The study showed that hand gestures aided in the emotion recognition rate for postures which others [Coulson 2004] had previously assumed as unimportant. Additionally, it was discovered that the emotion recognition rate using gestures can be greatly improved when emblematic actions are combined with functional actions. This study also confirmed our assumption that an integrated system covering facial expression, gesture and body postures is capable of enhancing the emotion recognition rate beyond the rates of the single systems.

INTRODUCTION

With the increasing prevalence of computers and other related technologies in many facets of today’s society, it becomes increasingly important to create systems capable of interacting well with humans. Non-verbal communication may be used to enhance verbal communication or even provide developers with an alternative for communicating information. Previous research has shown that emotions can be effectively portrayed through non-verbal means [Atkinson et al. 2004; Coulson 2004; Ekman 2003]. Facial expressions, posture, and gestures in particular have been recognized as an important modality for non-verbal communication and enable us to determine an individual’s mental and emotional state as well as his or her attitude/character traits (not covered in this study as the focus is on emotions).

Unfortunately, the audience’s recognition of these patterns is not universal and sometimes dependent on factors such as gender, cultural background, age of the individual, etc. To further complicate matters, humans are also capable of complex emotion displays such as portraying multiple emotions simultaneously through a process known as emotion blending, or to hide the perceived emotions from the observer by masking it with another emotion.

Thus, for the development of a virtual agent to effectively and accurately communicate to the user via non-verbal communications, we first need to study the mechanics behind emotion recognition before looking at methods for improving the emotion display of the virtual agent. The following section starts off with an overview of previous work.

FUNDAMENTALS OF EMOTION EXPRESSION

This section looks at how cues are able to produce emotions. In addition, it provides an overview of how certain factors may affect emotion recognition in a 3D virtual agent.

Introduction to Cues

Cues are non-verbal signals involving either the movement/positioning of individualized parts of the body or the movement/positioning of a group of body parts in concert with each other [Ekman 1978]. Cues such as a raised fist and narrowing of the eyes can indicate that the individual is angry. Movement/positioning classes like facial expressions, body postures, and gestures can involve one or more of these cues, which people (subconsciously) use for interpreting the emotional state.

Facial Expressions

Facial expressions involve facial cues that are displayed using body parts from the head region (e.g., eyebrows, mouth, lips). Common facial expressions such as the raising of the lips (facial cue) as part of a smile (facial expression) is interpreted by others to be a display of emotion of the actor; happiness in this example [Ekman 1978].

Body Posture

Body cues involved in body postures are displayed using body parts such as the torso, arms and legs. They are another component of non-verbal emotion expression. For example, the clenching of a fist and raising it to appear like the actor is trying to attack someone is usually interpreted by others as a display of anger [Ekman 2003].

Gestures and Actions

Gestures are actions/movements of body parts and they are another component of non-verbal communication of emotion. For example, a high frequency gesture such as jumping up and down very quickly can be interpreted by others to be a sign of happiness [Raouzaiou et al. 2004].

Factors Affecting Emotion Recognition

Although the exact factors which can influence the interpretation of emotion have not yet been thoroughly researched upon, four factors have recently surfaced based on current experiments and research. They are gender, job status, culture, and age.

Men and women express emotions differently [Brody and Hall 1992] in terms of the frequencies of occurrence (men often experience anger more often than women). It was also proven that recognition of ambiguous facial expressions is influenced by the gender of the person performing it [Condry 1976; Devine et al. 2000] whereby “masculine” emotions (e.g., anger) are assigned to men while “feminine” emotions (e.g., happiness) are assigned to women. As such, there is a need to be mindful of these gender stereotypes when trying to synthesize emotions.

Stereotypes of job status are known to exist too [Algoe et al. 2000]. For example, managers are often associated with “masculine” emotions and character traits while nurses are associated with “feminine” emotions and character traits. If the virtual agent is assigned human jobs (usually identified by the type of uniform they are wearing), ambiguous emotion expressions may lead others to wrongly assign “masculine” and “feminine” emotions to it.

Culture can also affect the interpretation of emotions [Bianchi-Berthouze el al. 2006]. It was discovered that the Japanese are less animated in their body expressions for emotion than the Sri Lankans leading to the same emotion being read differently.

Lastly, there is neurological evidence to suggest that age can affect the interpretation of emotions. It was shown that people in the 60-80 years old age group tend to suffer from emotion processing impairments and therefore require stronger or more cues to be displayed before being able to associate an emotion.

Emotion Blending and Transition

Human beings are capable of feeling multiple emotions simultaneously [Ekman 2003]. These emotions may transition/morph in time from one state of emotion to another (e.g., a loud noise may suddenly cause a passerby to feel surprise momentarily, which might later transition into a feeling of fear if the passerby feels that his/her life is in eminent danger), or they may also be displayed at the same point in time (e.g., the loss of a loved one in a car accident may cause a person to feel both angry and upset at the same time).

Emotion blending is the mechanism by which multiple emotional expressions are altered or combined simultaneously to convey more subtle information about the performer. Unfortunately, such a process is a complicated one and has not yet been well understood and researched by behavioral psychologists and animators.

CONCEPTS FOR REPRESENTING AND MODELING FACIAL EXPRESSIONS, POSTURES, AND GESTURES

This section looks at the approaches taken by others to represent and model emotion expression. It also describes the approach taken by this study.

Representing and Modeling Facial Expressions

There are several methods for modeling facial expressions such that they can convey emotions in virtual agent. This range from using mathematical models such as muscle vectors, or mesh-spring setups, to using physiological models such as Ekman’s facial action coding system (FACS) and the MPEG-II coding for facial action units (FAPS).

The approach taken by this study utilizes FACS [Ekman et al. 1978] as it has been around since 1978 and has undergone several rounds of testing and fine-tuning by the community. In addition, FACS is relatively easy to implement as it only considers the visible facial changes that result from muscle activity rather than simulating the entire muscle itself. This means that superimposing each facial change (known as an AU[1] in FACS) would be possible as each AU is unique and does not interfere with each other.

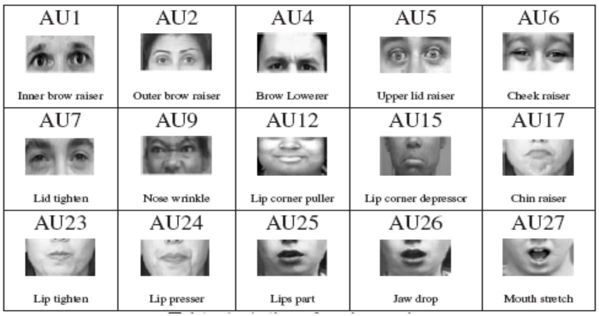


Table 1a: Some AU and their associated facial change obtained from Ekman’s study [Ekman 1978].

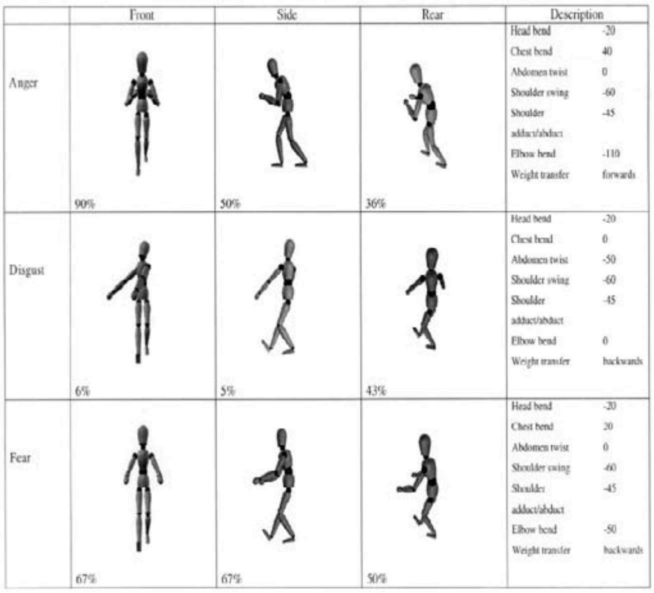
Basic Expressions	Involved Action Units
Surprise	AU1, 2, 5, 15, 16, 20, 26
Fear	AU1, 2, 4, 5, 15, 20, 26
Disgust	AU2, 4, 9, 15, 17
Anger	AU2, 4, 7, 9, 10, 20, 26
Happiness	AU1, 6, 12, 14
Sadness	AU1, 4, 15, 23

Table 1b: Table of the six basic emotions and the AUs involved.

(1) AU denotes an 'Action Unit' and each AU represents a unique facial change arising from a single muscular activity.

Representing and Modeling Postures

There exist a variety of sources which offer more or less detailed descriptions of emotional postures [Birdwhistell 1975; Boone and Cunningham 2001; Darwin 1872].
For instance, in the descriptions put forward by these authors, anger is described as involving a jutting chin, angular body shape, forward weight transfer, chest out and angled forwards, and a bowed head. Unfortunately, prior to Coulson's study [Coulson 2004], no formal research has been done to quantify the anatomical features which produce the emotional posture (i.e., posture was mostly descriptively documented).
Coulson found that the anatomical descriptions obtained from the studies mentioned earlier could be translated into joint rotations, which he then attempted to quantify via testing on human volunteers.
The approach taken by this study to modeling postures relies on Coulson's findings for the angle representation of each emotion as it is the only study which quantifies the respective joint angles.



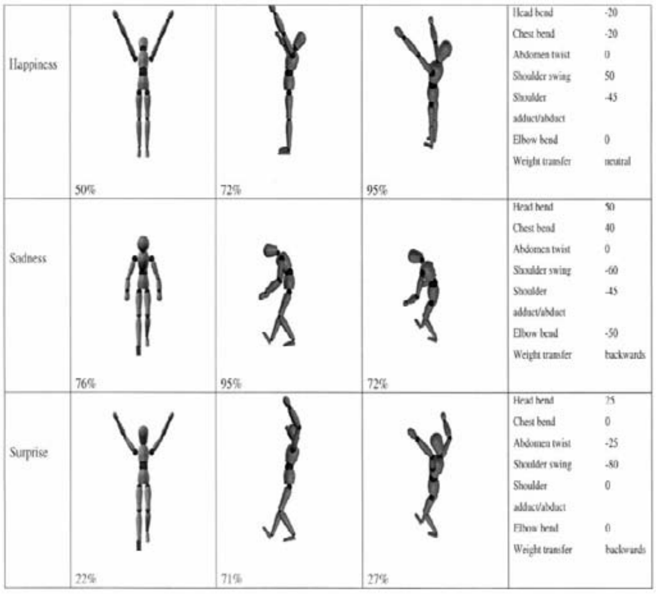


Table 2: Joint angles for the 6 basic emotions obtained from Coulson's study [Coulson 2004].

Representing and Modeling Gestures

Gestures in this study are modeled by animating the virtual agent since gestures are essentially non-static displays of emotion. As there is a relative paucity of studies on dynamic emotion gestures [Atkinson et al. 2004], the approach taken by this study relies on actors' knowledge of gestures. Raouzaiou et al. (2004) and Atkinson et al. (2004) formulated a short table of emotions that depict general hand and head gestures for each emotion.
This study relies upon those data to provide a basic framework for generating gestures as it is also compatible with the joint-angle system used for modeling of posture.

Emotion	Gesture Class
Joy	*hand clapping-high frequency*
Sadness	*hands over the head-posture*
Anger	*lift of the hand- high speed*
	italianate gestures
Fear	*hands over the head-gesture*
	italianate gestures
Disgust	*lift of the hand- low speed*
	hand clapping-low frequency
Surprise	*hands over the head-gesture*

Table 3: Table of gestures extracted from Raouzaiou's study [Raouzaiou et al. 2004].

SYSTEM INTEGRATION

Previous studies have focused on modeling facial expressions, body postures, or gestures on their own but few studies (if any) have been conducted on an integrated approach whereby all three expressions are modeled at the same time.
To determine the efficacy of an integrated system, three rounds of tests involving human volunteers were carried out. The tests were conducted over the internet in the form of a specially designed internet survey with images and animations embedded into the webpage and a simple short

question asking the participants to select the emotion that is best portrayed in the images/animations.

Each round of testing (except for the preliminary round) involved three sections that the volunteers have to complete. In the first section, the volunteers were presented with images of only the virtual agent's face expressing different facial expressions. The virtual agent's full body (including the face) was then made visible in the next section, where the virtual agent was expressing emotion through a combination of body postures and facial expressions. Finally, in the last section, the volunteers were presented with an animated video clip showing the virtual agent performing emotions through a combination of gestures, postures, and facial expressions.

Preliminary Testing

A round of preliminary testing was first conducted among six close acquaintances to stress-test the system for bugs and to optimize the survey pages. Additionally, they were also asked to determine the facial expressions displayed by the virtual agent.

The results obtained for the facial expression recognition are shown in Table 4. It was noted that anger and happy were already capable of obtaining a 100% recognition rate. Sad obtained an 83% recognition rate. On the other hand, surprise, fear and disgust were not capable of achieving any significant recognition rate (below the probability of a random guess).

Emotion	No. of correct recognition	Emotion Recognition Rate
Happy	6	100.00%
Anger	6	100.00%
Sad	5	83.33%
Surprise	3	50.00%
Fear	2	33.33%
Disgust	0	0.00%

Table 4: Facial expression recognition rate from preliminary testing.

First Round of Testing

After the preliminary round, the survey was officially launched and a total of 16 fully completed responses were collected from unique individuals. The results collected for facial expressions are shown in Table 5. The recognition rate for surprise and fear increased to more significant figures possibly due to the larger sample size for this round of testing.

Emotion (Facial Expressions)	No. of correct recognition	Emotion Recognition Rate
Happy	16	100.00%
Anger	16	100.00%
Sad	13	81.25%
Surprise	10	62.50%
Fear	11	68.75%
Disgust	2	12.50%

Table 5: Recognition rate from facial expressions alone.

It was also observed that the facial expression for surprise (10/16) and fear (11/16) was commonly mistaken for one another. This could possibly be due to the fact that both emotions share some similarities such as the opened and laterally stretched mouth and the raised eyebrows (they both have AU1, AU2, AU5, AU15, AU20, and AU26. Surprise only differs from fear by 2 AUs – AU4, AU16). It was postulated that the confusion between the two facial expressions could perhaps be resolved by creating a bigger difference in appearance of the mouth region (making the mouth open wider to show more teeth for fear expression). This modification was applied to the second round of testing.

Figure 1: Facial expression of surprise (left) and fear (right)

When posture was added to the respective facial expression of the emotion, the results obtained are summarized in Table 6.

Emotion (Postures)	No. of correct recognition	Emotion Recognition Rate
Happy	16	100.00%
Anger	16	100.00%
Sad	16	100.00%
Surprise	9	56.25%
Fear	14	87.50%
Disgust	1	6.25%

Table 6: Recognition rate from integrating posture and facial expression.

Fear (14/16) now fared 18.75% better than in the facial expressions test (11/16) while sad was now capable of obtaining a 100% recognition rate.

Surprise (9/16) fared almost similarly as in the facial expressions test while anger (16/16) and happy (16/16) once again achieved 100% recognition rate.

From this test, it appears that integrating posture with facial expression was capable of improving the recognition rate for some emotions (fear and sad) without any adverse effects to the other emotions (surprise, happy, anger, disgust).

In addition, it was once again noted that surprise was sometimes mistaken for fear and this could likely be due to the hands up posture in surprise closely resembling a surrendering position that is commonly associated with fear as well (Figure 2, left).

It was speculated that the occasional confusion between surprise and fear could possibly be resolved by modifying the hands up posture of surprise in a more distinguishing manner whereby the palms are facing outwards and the fingers are slightly curled which reduces the appearance of a surrendering stance that is frequently associated with fear (Figure 2, right). This new posture was subsequently used for the second round of testing.

Figure 2: Surprise posture - old (left) and new (right)

When gesture was integrated with the respective posture and facial expression, the results obtained are shown in Table 7. A website has also been created for viewing[2] the animated gestures as well as for downloading[3] them.

The recognition rate of happy, anger, and sad remained at 100% while fear and disgust obtained approximately similar recognition rates as during the posture test.

(2) http://emotionexpression.yolasite.com
(3) http://www.filefactory.com/file/ag65dgd/n/artwork_pptx

The greatest disparity came from surprise (25%), which had its recognition rate halved as compared to the posture test (56.25%). One possible explanation was that the gesture for surprise was unrealistic as it required the model to lean backwards while lifting its arms. This created an appearance of a falling motion and could have perhaps confused the test participants.

Emotion (Gestures)	No. of correct recognition	Emotion Recognition Rate
Happy	16	100.00%
Anger	16	100.00%
Sad	16	100.00%
Surprise	4	25.00%
Fear	12	75.00%
Disgust	2	12.50%

Table 7: Recognition rate from integrating gesture, posture, and facial expression.

Disgust (Figure 3) was unable to achieve any significant recognition rate in all 3 sections of the test. A possible explanation could be that disgust is rarely associated with many expressions thus leading to it being more difficult to identify.

Figure 3: Disgust facial expression (left) and posture integrated with facial expression (right)

Independent studies by Coulson [Coulson 2004], Hofman and Walk [Hofman and Walk 1984] have similarly concluded that disgust was usually the poorest performing emotion of the six basic emotions. As such, extra attention was set aside to overhaul this emotion by collecting data for the disgust gesture through observing videos of real human emotions. Our initial attempt at analyzing video corpus was not successful as these video sources were often limited in supply and were not easily attainable. An innovative approach was instead carried out by observing hours of unscripted reality TV shows such as "Just for Laugh Gags" and "The Price is Right" as these were frequently broadcasted on public TV network and the emotions portrayed by the subjects tend to be more natural and realistic than those of scripted TV shows. The new posture and gesture that we obtained for disgust no longer followed Coulson's joint angles data as this new posture now involves covering the mouth/nose with the palms of the hand as though shielding it from something repulsive (Figure 4).

On the whole, integrating more than 1 type of expression appeared to have either a positive or neutral effect on the recognition rate for most emotions except for surprise which resulted in a negative impact. It was also observed that as more and more types of expressions are integrated together, the recognition rate appeared to improve further.

Figure 4: Disgust posture - old (left) and new (right)

Second Round of Testing

The changes proposed earlier were applied to this new round of testing. A total of 25 fully completed responses from unique individuals were collected.

For facial expressions, the results obtained are shown in Table 8.

Emotion (Facial Expressions)	No. of correct recognition	Emotion Recognition Rate
Happy	25	100.00%
Anger	25	100.00%
Sad	24	96.00%
Surprise	17	68.00%
Fear	16	64.00%
Disgust	4	16.00%

Table 8: Recognition rate from facial expressions alone.

A minor tweak was made to the sad facial expression by modifying the mouth to be more upturned/inverted appearance to increase the intensity of the sadness portrayed. That seemed to have increased the recognition rate from 81.25% to 96%. This implies that the intensity of the emotion can adversely affect the emotion recognition rate and if one is not careful about the changes made to certain facial regions, one may unknowingly alter the emotion's recognition rate.

The attempt at modifying the mouth region to aid in distinguishing the facial expression of fear from surprise appeared to have no effect on the recognition rate of the two emotions. This could possibly imply that the two emotions are not distinguished by the mouth as was earlier postulated during the first round of testing.

For the other emotions, their facial expression recognition rate remained approximately similar which was expected since no modifications was made to them.

Figure 5 summarizes the results obtained from the three rounds of facial expressions testing.

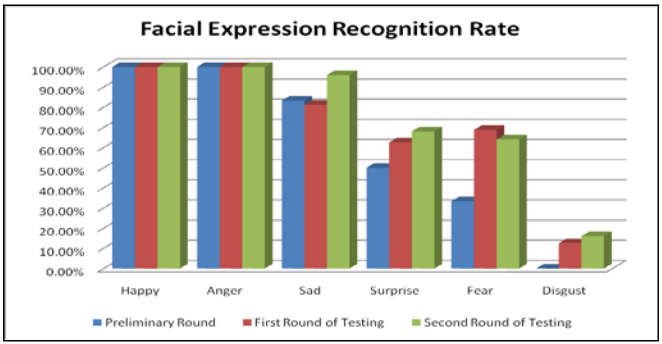


Figure 5: Facial expression recognition rate for the 3 rounds of testing

When posture was added to the respective facial expression, the results obtained are summarized in Table 9.

Emotion (Postures)	No. of correct recognition	Emotion Recognition Rate
Happy	25	100.00%
Anger	25	100.00%
Sad	25	100.00%
Surprise	17	68.00%
Fear	25	100.00%
Disgust	7	28.00%

Table 9: Recognition rate from integrating posture and facial expression.

Happy, anger, and sad maintained a 100% recognition rate which was similar to the results for the earlier round of testing.

The confusion between surprise and fear appeared to have been resolved as evidence by the recognition rate for surprise increasing by almost 12% after making the earlier mentioned changes to the hand posture. This implied that hands do appear to play a part in emotion expression too. Coulson had ignored the posture of the hands in his study as he felt that they were not capable of conveying any cues but the test conducted in this study showed that the positioning of the hands could be used as markers for distinguishing between fear and surprise.

Disgust showed the biggest improvement in this round of testing (from an initial recognition rate of 6.25% to 28%) after it was completely reworked upon. A likely explanation for this improvement is that Coulson's approach to posture expression relied solely on functional actions whereas the recent changes made in this study added in an additional emblematic action component. Functional actions refer to the behavioral significance of an emotion. They are generally determined by the tenseness of the muscle due to an emotion, which subsequently results in a particular posture being presented. Coulson's study focused on documenting that tenseness in the form of joint angles. Emblematic actions on the other hand refer to the gestures that are commonly associated with the emotion.

The functional action for disgust is to turn away from the source of repulsiveness whereas the emblematic action is to raise the hands to cover the nose/mouth as though trying to shield off those crevices from something toxic. By employing this additional emblematic approach, new cues can be added to the posture expression of an emotion. The tests conducted in this study showed that it is helpful in overcoming the recognition problems of weak emotions such as disgust as these emotions usually have very few distinct cues that people recognize.

Figure 6 summarizes the results obtained from the two rounds which involved integrating facial expressions with body posture.

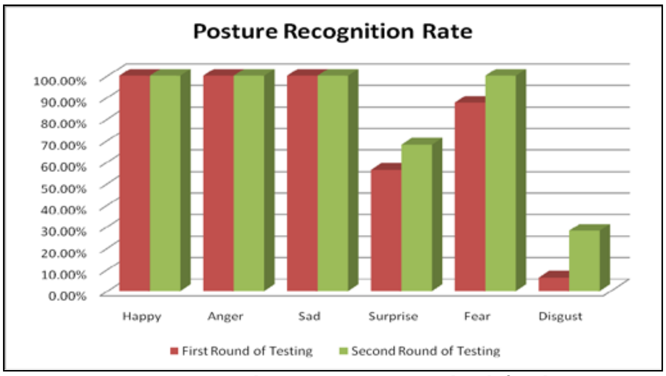


Figure 6: Recognition rate when posture is integrated with facial expression.

When gesture was integrated with the respective posture and facial expression, the results obtained are shown in Table 10.

Emotion (Gestures)	No. of correct recognition	Emotion Recognition Rate
Happy	25	100.00%
Anger	25	100.00%
Sad	25	100.00%
Surprise	20	80.00%
Fear	23	92.00%
Disgust	17	68.00%

Table 10: Recognition rate from integrating gesture, posture, and facial expression.

Happy, anger, and sad once again obtained a 100% recognition rate which was similar to the previous round of testing.

Surprise, fear and disgust on the other hand had their recognition rates significantly improved.

Surprise's recognition rate increased from 25% in the previous round of testing to 80% after modifying the unrealistic bending animation. The new positioning of the palm and the curling of the fingers that was adopted for the posture expression might have also contributed to this improvement.

Fear's recognition rate increased from 75% to 92%, which was likely due to the minor tweak that was done by having the virtual agent turn its head away from the source of fear. Once again, this proves that altering the intensity of an emotion can affect its recognition rate.

Disgust also exhibited a huge improvement in its recognition rate (from 12.5% to 68%) after its gesture was overhauled to include emblematic actions. This proves the importance of having emblematic actions to aid weak emotions.

Figure 7 summarizes the results obtained from integrating gesture with body posture and facial expression.

By comparing Figures 5-7, it was observed that integrating gestures with body postures and facial expressions produced more superior results than relying only on a combination of body posture and facial expressions, and this is in turn more superior than solely relying on facial expressions alone.

In addition, the increment in recognition rate from the integration of the various types of expressions was non-linear and was capable of increasing exponentially for emotions such as disgust.

The caveat emptor for implementing integrated systems is that extra care must be taken to ensure that the system is optimally tweaked or it might result in a lower emotion recognition rate; such as in the case of surprise during the first round of testing whereby an unrealistic surprise gesture lowered the recognition rate instead of improving it.

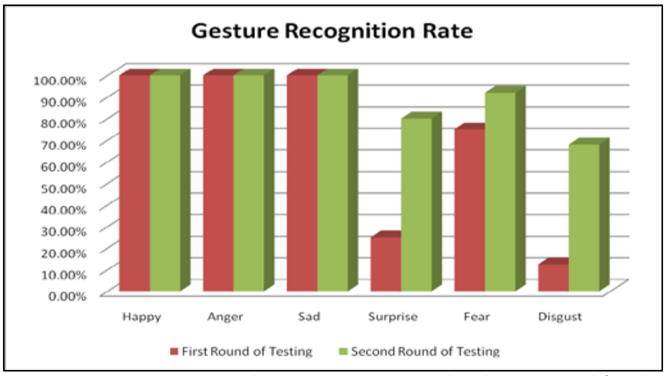


Figure 7: Recognition rate when gesture is integrated with posture and facial expression.

CONCLUSION

This study looked at the individual systems for synthesizing facial expressions, body postures, and gestures and the issues associated with these systems when they are implemented in a virtual agent. It then investigated the effects of integrating the various systems together before making improvements to the individual systems to enable them to work in synergy with one another.

The tests that were carried out revealed that an integrated system of gestures, body postures, and facial expressions that is properly tweaked was capable of producing better emotion recognition rates than a system which only relied on body postures and/or facial expressions.

This study is also an extension of Coulson's research as it shows how the posture and gesture of weak emotions such as disgust can be improved upon by relying on additional emblematic actions on top of functional ones used by Coulson in his study.

Lastly, this study showed that Coulson was wrong in his assumption that hand gestures were not important for emotion recognition. The tests conducted in this study revealed that hand gestures may serve as markers for differentiating between certain emotions such as fear and surprise. Based on facial expressions and body posture, fear and surprise appear to be quite similar, often confusing the test volunteers.

Future Directions

Future directions for improving this study are to look at methods for implementing advanced emotion display in virtual agents to enable them to be capable of emotion masking, emotion blending, and emotion transition.

Emotion masking allows the virtual agent to display a different emotion from what it currently perceives or experiences [Buisine et al. 2006]. People usually perform this due to factors such as cultural backgrounds, gender, etc. (apart from conscious manipulation). For example, public figures often have to hide their emotions – especially those of sadness – as showing sadness is often regarded as a sign of weakness in many cultures.

The same can be applied to virtual agents in computer games. For example, the heroic character in a computer game often has to appear brave and courageous at all times. This can be achieved by masking negative emotions such as sadness and fear even when the character is outnumbered by enemies and is facing impossible odds of winning. In doing so, the character appears more human-like to the user since the user is able to sense that the character is merely putting up a front to keep up with its appearance while deep down, it is still experiencing human emotions (i.e., fear). If emotion masking was not applied here, the character would appear to be unfazed by the precarious situation that it is in and this can lead to a reduction in the human-like realism and emotional depth of the character.

Another potential area of research is emotion blending. Emotion blending allows multiple emotions to be combined and displayed simultaneously. Current studies [Cowie et al. 2002; Noot et al. 2003; Albrecht et al. 2005; Niewiadomski et al. 2005; Buisine et al. 2006] are limited to the blending of only two facial expressions simultaneously as they rely on the concept of allocating a positive emotion (e.g., happy) to the lower half of the face while reserving the upper half for negative emotions (e.g., anger). This concept was based on Gouta's [Gouta and Miyamoto 2000] and Bassili's [Bassili 1979] research. However, there have been no formal studies done to date on performing emotion blending with body postures and gestures.

If the face and body could be further subdivided into more than two regions, it might be possible to blend multiple emotions simultaneously. An allocation system could perhaps be worked out to determine which region a particular emotion is most commonly recognized from. That emotion would then have priority over that region of the face or body. This may enable the blending of multiple emotions while at the same time ensuring minimal impact to the recognition rates of the various emotions being combined since the more important cues of each emotion are preserved. This method may also make it possible to blend body postures and gestures.

Emotion blending is important for creating realistic virtual characters because humans are capable of feeling and expressing multiple emotions at the same time. For example, when a character in a computer game witnesses a close friend being slain by an evil villain, the character might be programmed to feel sadness (from losing a friend) and anger (towards the villain), perhaps even vowing revenge, which closely mimics the emotions felt by real people in similar situations. An accurate blending of these two emotions would thus be essential to allow the player to associate more closely with the character since the player is able to comprehend what the character is going through.

Another area of research to look at is emotion transitioning. It describes how the various face and body parts change in time when an emotion changes to another emotion. This intermediate process usually occurs for a fraction of a second and is often used for detecting if someone is lying [Ekman 2003]. This could perhaps be employed by characters in computer games for scenarios where the character needs to tell a lie. If done improperly, the change in emotion could appear awkward and unrealistic and the user may not be able to pick out the subtle cues to indicate that the character is not truthful.

Lastly, it would be beneficial to look into the creation of a database for universally accepted gestures of each emotion. Currently, there is scarce literature on the specific gestures that are commonly associated with each emotion. A database of gestures could possibly be constructed by observing large quantities of video corpus (often limited in supply and not readily available) or by observing unscripted TV reality shows, which was the approach taken in this study. Alternatively, the gesture database could also be formulated based on actor's impression since they are more adept at portraying emotions intentionally for their audience.

ACKNOWLEDGMENTS

This work was supported by the Singapore National Research Foundation Interactive Digital Media R&D Program under research Grant NRF NRF2007IDM-IDM002-051, as well as by a joint grant of Singapore's Ministry of Education / Academic Research Fund and the National University of Singapore under research Grant RG-263-001-148.

REFERENCES

Albrecht, I. and Haber, J. and Schröder, M. and Seidel. 2005. "Mixed Feelings: Expression of Non-Basic Emotions in a muscle-based Talking Head." *Special issue of Journal of Virtual Reality on"Language, Speech & Gesture"*, no. 8(4): 201-212.

Algoe, S. and Buswell, B. and Delamater, J. 2000 "Gender and Job Status as Contextual Cues for the Interpretation of Facial Expression of Emotion." *Sex Roles: A Journal of Research*, no. 42 (Feb): 183-208.

Atkinson, P. and Dittrich, W.H. and Gemmell, A.J. and Young, A.W. 2004. "Emotion perception from dynamic and static body expressions in point-light and full-light displays." *Perception*, no. 33(6): 717–746

Bassili, J.N. 1979. "Emotion Recognition: The Role of Facial Movement and the Relative Importance of Upper and Lower Areas of the Face." *Journal of Personality and Social Psychology*, no. 37(11): 2049-2058.

Bianchi-Berthouze, N. and Kleinsmith, A. and Ravindra, P. 2006. "Cross-cultural Differences in Recognizing Affect from Body Posture." *Interacting with Computers*, no. 18 (Dec): 1371-1389.

Birdwhistell, R.L. 1975. "Background Considerations to the Study of the Body as a Medium of Expression." In *The Body as a Medium of Expression*, J. Benthall, eds. Suffolk, Allen Lane, 36-58.

Boone, R.T. and Cunningham, J.G. 2001. "Children's expression of emotional meaning in music through expressive body movement." *Journal of Nonverbal Behavior*, no. 25: 21-41.

Brody, R. and Hall, A. 1992. "Gender and Emotion." In *Handbook of Emotions*, M. Lewis and J.M. Haviland, eds. New York, Guilford Press, 447-460.

Buisine, S. and Devillers, L. and Martin, J., and Niewiadomski, R. and Pelachaud, C. 2006. "Multimodal Complex Emotions: Gesture Expressivity and Blended Facial Expressions." *International Journal of Humanoid Robotics, special issue Achieving Human-Like Qualities in Interactive Virtual and Physical Humanoids*, no. 3 (3): 269-291.

Condry, J. and Condry, S. 1976. "Sex Difference: A Study of the Eye of the Beholder." *Child Development*, no. 47 (Sep): 812–819.

Coulson, M. 2004. "Attributing Emotion to Static Body Postures: Recognition, Accuracy, Confusions, and Viewpoint Dependence." *Journal of Nonverbal Behavior*, no.28 (2): 117-139.

Cowie, R. and Douglas-Cowie, E. and Kollias, S. and Raouzaiou, A. and Tsapatsoulis, N. 2002. "Emotion Recognition and Synthesis based on MPEG-4 FAPs" In *MPEG-4 Facial Animation: The Standard, Implementation and Applications*, I.S. Pandzic, I.S and R. Forchheimer, eds. New York, USA, 141-167.

Darwin, C. 1872. *The Expression of the Emotions in Man and Animals*. University of Chicago Press.

Devine, G. and Hyde, S. and Keltner, D. and Plant, A. 2000. "The Gender Stereotyping of Emotion." *Psychology of Women Quarterly*, no. 24: 81-92.

Ekman, P. and Friesen, W.V. and Hager, J. 1978. *Facial Action Coding System – The Manual*. Research Nexus Division of Network Information Research Corporation.

Ekman, P. 2003. *Emotions Revealed*. Times Books, New York, N.Y.

Gouta, K. and Miyamoto, M. 2000. "Emotion Recognition, Facial Components Associated with Various Emotions." *Shinrigaku Kenkyu*, no. 71(3): 211-8.

Hofman, C. P. and Walk, R. D. 1984. "Emotion and Dance in Dynamic Light Displays." *Bulletin of the Psychonomic Society*, no. 22(5): 437–440.

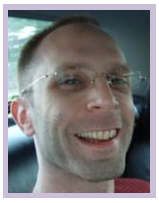
Niewiadomski, R. and Ochs, M. and Pelachaud, C. and Sadek, D. 2005. "Intelligent Expressions of Emotions." In *Proceedings of the 2005 1st International Conference on Affective Computing and Intelligent Interaction* (Beijing, China, Oct. 22-24). ACII, Beijing, China, 707-714.

Noot, H. and Ruttkay, Z. and Ten Hagen, P. 2003. "Emotion Disc and Emotion Squares: tools to explore the facial expression face." *Computer Graphics Forum*: 22(1): 49-53.

Raouzaiou, A. and Karpouzis, K. and Kollias, S. 2004. *Emotional Synthesis in Virtual Environments*. National Technical University of Athens.

BIOGRAPHY

Shawna Tan is a final year undergraduate from the Department of Electrical & Computer Engineering at the National University of Singapore. She has a profound interest in electronic gaming particularly in the area of game related artwork & animations, as well as in game design. She has recently completed her final year thesis on the topic of non-verbal emotion expressions for virtual agents.



Alexander Nareyek is the director of the interactive intelligence Labs (ii Labs) and of the multi-disciplinary Games Lab at the National University of Singapore, where he also holds an assistant professorship. He received his diploma and Ph.D. from the TU Berlin/Germany, held positions at GMD-FIRST/Germany, Carnegie Mellon University/USA, and the Cork Constraint Computation Centre/Ireland, and served as CEO and CTO for Digital Drama Studios/Czech Republic. He is the founder and owner of the games company Kingdom Crafting Pte Ltd, and also serves as chairman of the International Game Developers Association's Artificial Intelligence Special Interest Group.

ATTENTION DEFICIT DISORDERS WITH HYPERACTIVITY TREATMENT USING GAMES AND WIIMOTE

Begoña García, Amaia Méndez, and Ibon Ruiz
University of Deusto
Avda. Universidades 24, 48007
Bilbao, Spain
E-mail: {mbgarcia, amendez, ibruiz}@eside.deusto.es

KEYWORDS
Wiimote, ADDH, leisure time, learning, user interaction

ABSTRACT

This paper deals with the design and development of a multimedia game for the analysis and treatment of Attention Deficit Disorders with Hyperactivity (ADDH).
The treatment has been planned at three levels. Each level treats a characteristic related to socialisation, integration and the expression of feelings. For this purpose, various scenarios have been described (the city, the college and the home) and the user has to pass the test related to social situations.
To interact and play this game, people will use the Wiimote, chosen due to its simplicity and usability. The results produced by the game are saved in a database to be analysed by the instructors at a later date.
The game contains different situations (leisure time, learning, physiotherapy) to work on the social abilities of the user. The validation of this new therapy will be carried out with the help of neuropsychology experts (from the University of Deusto), children and patients with this disorder.

1. INTRODUCTION

The world of games may have countless possibilities for application. Although one of the main aims tends to be that of occupying leisure time and providing entertainment, in the case of the project presented in this paper, the direct application of the game [1] is of interest to the rehabilitation, treatment and education of children suffering from the neuro-development disorder called Attention Deficit Disorders with Hyperactivity (ADDH).

ADDH refers to a chronic disorder that initially manifests itself in childhood and is characterized by hyperactivity, impulsivity and/or inattention. Not all of those affected by ADDH manifest all three behavioural categories. These symptoms can lead to difficulty in academic, emotional and social functioning. The diagnosis is established by satisfying specific criteria and may be associated with other neurological, significant behavioural and/or developmental/learning disabilities. Therapy may consider the use of medication, behavioural therapy, and adjustments in day-to-day lifestyle activities [2].
So as to be able to tackle this project, a multidisciplinary team made up of computer engineers, telecommunications engineers, social educators and neuro-psychologists provided their help. The latter’s role consisted of defining the specific needs of the ADDH children collective and that of proposing a coordinated form of treatment alongside the engineers, in such a way that a technical game was adapted to each idea from the psychological sphere; moreover, the game can be objectively assessed through reports. The collaboration of the Association for People with ADDH was vital for the realization of experimental trials with children suffering from this pathology.

2. OBJECTIVES

The overall aim agreed on for this Project is to produce an interactive game for children with ADDH, one that enables them to develop social relations and how to express their feelings, based on their living real situations on a PC which they complete in the following situations: home, school and the town. In order to pass from one level to the next, the program will assess and tutor the degree of development achieved in the aspects dealt with.
This general aim can be broken down into a group of specific aims, as follows:

- Medical Objective: to treat and rehabilitate children with ADDH with the aid of a user-friendly, simple and attractive tool.
- Technological Objective: to design and develop an accessible computer application that implements games with semi-real characters whose image corresponds to that of an ADDH patient. It also has to allow interaction by means of the Wii remote control stick.
- Social Objective: to aid the integration and socialisation of ADDH children in the home, at school as well as in urban environments.

It was this group of aims, involving not only technological experts but also doctors, psychologists and associations, that made it possible to carry out this successful pilot experience with a charming collective of children with difficulties.

3. METHODOLOGY

In this section, the authors describe the methods used to develop the game and the interaction with the Wiimote. For the selection of the suitable 3D design software to be employed for the modelling of the game's various components, the following were considered: Autodesk 8.5, Luxology Modo 301, Newtek Lightwave 9.3 and Softimage XSI Foundation 6.2. Such parameters as the platforms they accept, the documentation they provide, that they are easy to use, etc., were taken into account, as can be seen in Table 1.

	Autodesk Maya Completo 8.5	Luxology Modo 301	Newtek Lightwave 9.3	Softimage XSI Foundation 6.2
Platforms				
PLE	Yes	None	None	None
For Industries:	Very good	Low	Very good	Good
Interface	Flexible and Intuitive	Excelent	Texts, interface a bit dated	Logical Text and friendly interface
Documentation	Excelent	Very Good	Excelent	Very Good
Quality	Excelent	Excelent	Excelent	Excelent
Quality/Plugin	Excelent	NA	Excelent	Excelent
Textures	Very good	Excelent	Excelent	Very good

Table 1. Comparison between the 3D modelling software

In accordance with the characteristics analysed, the software chosen was: Autodesk Maya 8.5, because it is Multiplatform.
The Wiimote control stick was chosen for interaction with the ADDH user as it is regarded as by the Spanish Special Network Federation (SSNF) as being a novel means of interaction.

4. DESIGN

The game developed comprises 3 scenarios: at home, at school and in town. A photo of the user is taken with a webcam and, from that moment on, this will be the face of the main character in the game so that the user feels more identified with the situation.
The design was made in a scalable and modular fashion, so that it can easily be amplified and modified, even by non-specialists.

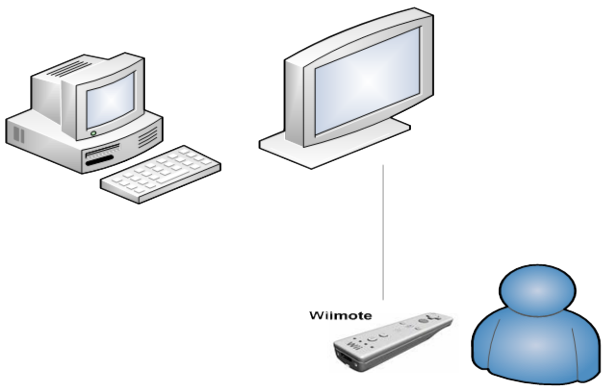


Figure 1. Interaction with the Wiimote

The Wiimote (the interaction can be seen in figure 1) has 11 programmable keys, movement sensors, Infrared and Bluetooth for communication, which in our case is with a PC (using C# programming language) as the human interface device (HID). Within the context that concerns us, that of ADDH children, the 11 keys are not used, but just the larger ones, which are easier to use and also intuitive; the functions of the other keys will be deactivated.
Three levels, which the children will have to complete, have been developed for each scenario. Such abilities/skills as responsibility, confidence or Independence are worked on at each level.
In figure 2, the activities diagram for one of the options at the elementary level can be observed: the "asking for help" option is in operation here. Given that this is the elementary level and that the users have to get used to the game and the

Wiimote interaction, the number of options available to the user is quite limited.

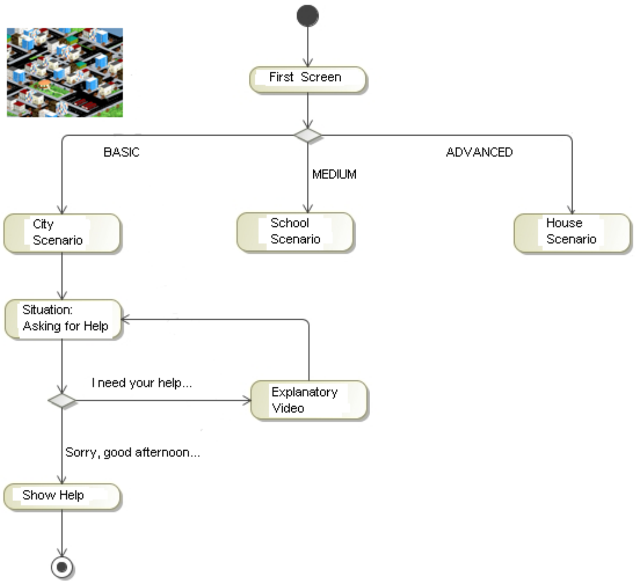


Figure 2. Activity Diagram

At the most advanced level, the user is faced with situations closer to reality, bringing together all the skills previously focused on at the elementary and intermediate levels. There are multiple options available for resolving each situation in all cases.
The user must always know if the chosen option is correct and why, for which purpose explanatory videos are employed. The user receives positive awards for work well done (medals, for example), and the progress made is logged. The instructors will revise the results in order to assess the validity of this new treatment.

5. RESULTS

In order to assess the results of this project, the objective fulfilment indicators put forward from the beginning have to be taken into account.
We will therefore go over the above mentioned aims, providing the results obtained for each one.
As regards the medical objective of treating and rehabilitating ADDH children with the aid of a friendly, simple and attractive tool, a satisfaction questionnaire was filled in by the professionals in neuropsychology in charge of evaluating the application developed.
As can be appreciated in table 2, the game proved satisfactory for experts in its treatment of the pathology, although they would carry out some modifications.

QUESTION	ASSESSMENT (0-Very poor/No; 5-Very good/Yes)
Are you satisfied with the "game" developed?	4
Do you feel this therapy can improve on traditional methods?	3
Do you think the treatment levels are suitable?	5
Would you make changes to the home, school and town scenarios?	3
Are the progress reports suitable?	4
Would you make any changes to the program?	5

Table 2. -Satisfaction questionnaire for professionals diagnosing and treating ADDH

Since the comprehensive treatment, during which the child passes through the 3 levels, takes a considerable length of time, we do not yet possess objective data in terms of final results enabling us to determine the difference between using this game and following traditional methods.
One of the suggestions made by the neuropsychologists to the technicians involved in developing the game is that the third scenario, corresponding to the integration and realisation of activities in the town, should come as close as possible to the patient's situation. This would mean introducing the simulation of a town (see Figure 3) that the child could relate to, it being similar to where he actually lives.

Figure 3. Example of the town scenario forming part of the third learning level

As regards the technological objective, we will see the results of the design and development of an accessible computer application which implements some games with semi-real characters, the appearance of whom corresponds to that of the ADDH patient. It must permit interaction by means of the Wii control stick. The generic person in Figure 4 has been designed so that the ADDH child can identify him/herself with the main character in the game. A real image of the child's face will be added at the beginning of the therapy.

Figure 4. Screen display of the game's main character representing the child user

The three planned scenarios have been developed: home, school and town. Figure 5 represents an example of the first learning level, which takes place in a fictitious home where the character carries out several activities favouring the ability to express feelings and wishes in the correct way.

Figure 5. Example of a household activity

To conclude this reflection upon the results of the project, assessment of the social objective remains. This aim intends to aid integration and socialization of ADDH children in the home, school and town environments. Two elements will be necessary to assess this feature. On the one hand, the patients under study need to have successfully passed the game's 3 levels. On the other hand, at this moment the suitability of the treatment should be assessed by the psychologists, as well as the day-to-day satisfaction obtained by the families with the new experiences they have had with their child following the therapy. As this will require at least a year of trialling, we have no objective data as yet. However, we certainly should take into consideration the enthusiasm, hope and confidence expressed by the relatives of children with ADDH who are testing the system.

6. CONCLUSIONS

As a global evaluation of the project, we can state that the technological innovation in the treatment techniques of ADDH has been well received in the professional sphere, by official associations and by the families of children with ADDH.

From a medical point of view, it would be ideal to have a personalised game for each child, developing a town scenario similar to where the child lives, for example. This is technologically possible, but demands an excessively individualized form of treatment for each patient

At the present time, trials are still being carried out with children diagnosed with ADDH, and the teams of neuropsychologists and patient associations involved are enthusiastic about the novel form of treatment. We hope that over the coming year we will obtain further objective data that will validate the suitability of the game for the treatment of not only this pathology but for others as well. In particular, if the results are suitable, a similar game for the treatment of infants affected by Asperger syndrome will be considered.

7. Acknowledgements

The authors of this paper would like to thank the students Nuria Aresti and Andoni Crespo for their help. It is important to remark the support given by the neuropshycology group of the University of Deusto.

REFERENCES

[1] P. Chan and T. Rabinowitz. "A cross-sectional analysis of video games and attention deficit hyperactivity disorder symptoms in adolescents". Ann Gen Psychiatry. 2006; 5: 16.

[2] Attention Deficit Disorders with Hyperactivity (ADDH). National Institute of Mental Health. U.S. Department of HealthandHuman Services.

Google: Games and Gaming, 2009 to 2015

Stephen E. Arnold

Arnold IT
Harrod's Creek, Kentucky 40027, USA
sa@arnoldit.com

Introduction

I want to provide a quick review of Google's approach to games and gaming. I want to show some screenshots that make clear that simple and more complex games are available with more games becoming available everyday. I then want to describe how Google views the notional topic of games for users of computing devices. I want to conclude by putting my remarks in a timeline that carries the subject of Google and games to the year 2015. The date is not arbitrary because Google works in chunks of three to five years. Google's approach to games won't change too much in the next 16 years, but the scope, application, and monetization of games and game technology will. My conclusion may surprise you. By 2015, Google may be one of the leading game platforms with a broad range of products and services that use gaming technology in interesting, revenue-boosting ways.

Google People

Most users of Google's systems don't know individual Google engineers by name. The company has nearly 19,000 professionals on staff and about two thirds of them are engineers, computer scientists, mathematicians, or physicists.

Quite a few people today play games. The devices range from the high-end, state-of-the-art platforms like the Microsoft Xbox, the Nintendo Wii, and the Sony PS3 to the grandma friendly games on Yahoo or mobile phones. One store in rural Kentucky where I live sells a $5 keychain with a simple game for bored adults and affluent seven year olds to play when stuck behind a horse in Harrod's Creek. New platforms bring new people to the "game party". Add in the influx of mobile device users, and the stage is set for a "game revolution".

I want to highlight two Google engineers and mention some of their work to give you an idea about how deep game technology has been embedded at Google.

Steve Lawrence, an Australian, is a gamer. In addition to authoring technical articles that have been referenced more than 5,000 times, he is the author of Game sports betting markets, Sandip Debnath, David M. Pennock, C. Lee Giles, Steve Lawrence, ACM Conference on Electronic Commerce, 2003, pp. 258-259. Dr. Lawrence is one of Google's most prolific inventors, and his technical skill has influenced inventions ranging from user interface (US7272601) to personalized network searching (US2008/0215553).

Consider Ross Koningstein. He was a graduate student at Stanford's Aerospace Robotics Laboratory when he contributed to the development of Chuck Yeager's Advanced Flight Trainer II and Car & Driver Text Track. Google's advertising system is based on bids. The methods used are dependent on calculations used in traditional games like horse racing. Mr. Koningstein has been working to bring game-like features and interface elements to Google's advertising management system. The idea is that a person at an ad agency can use a game-like method to model what certain types of ads and a specified amount of ad money will generate for the advertiser. Dr. Koningstein wants to make modeling ad spends and ad management more of an interactive game experience. You can read more about his approach in Google patent documents US20050228797, US20050096979, US20060224444, US20060224447, US20050114198, and several others.

You get the idea. Dr. Koningstein is not dabbling; he is inventing systems and methods that have roots deep in the interactive game experience.

Keep in mind that other Googlers have equally deep roots in gaming.

Google Technology

I want to do a quick fly through of Google technology and provide you with some screenshots of applications that are available today.

First, Google is a platform, and it offers a range of software development kits, application programming interfaces, and "sandbox" toys. The idea is that a developer with online basic programming skills can use the Google platform. At the other end of the spectrum, a professional developer or a company focused on game development can create applications that run on the Google platform.

Keep in mind that the platform is a one way street. This means that you can put code into Google but it can be difficult to repurpose that code for another platform. Therefore, the best way to think about using the platform for a game or some other application is to create a game for a platform such as the iPhone and then recycle the graphics and other useful bits for the Google Android platform. You will learn in a few moments that this recycling approach may be the path forward for the next few years.

Second, Google tried to cut a deal with Yahoo for online games in the 2005 to 2006 period. My sources suggested that the tie up did not make sense. Google on the surface has not played a major role in commercial game development. In fact, the model today is influenced by Google's need to be perceived as an open source company that is not a monopoly. The point is that if you get into the Google development space with a game, you will be operating in a competitive but open environment. At some point in the future, Google could change its approach, but there is little downside for experimenting with the Google platform. New tools such as Google Wave will be forthcoming. Coupled with Chrome (Google's virtual machine and container system) and Android (a chunk of the Google operating system), Google now offers a usable platform for game development.

Third, the forthcoming Google Wave technology (a component of Google's dataspace initiative) appears to be a significant new component of the Google service suite. The idea behind Wave is a plastic bag. Put carrots or small parts in the bag, and you can manage them. Wave allows a developer to create a space – a digital freezer bag. Activities can take place within the bag. Wave makes it possible to have the objects in the digital bag interact. The idea is to make it possible to create new types of social interactions with information objects. The most important feature is that states can be saved. It is possible to slice and dice the objects and the interactions by time. If you think about this functionality, new opportunities for games and game like experiences can be built on these multidimensional functions. Let me give one example: lectures, lab experiments, and student interaction. I think certain types of instructional constructs where traditional game like features and time can be combined with social interaction in useful ways.

Keep in mind that Google's technologies pivot on programming languages that many developers know. These include JavaScript, php, python, and Java, among others. The point is that you can hit the ground running with Google's sample code and your favorite programming language. At this time, there's no fee, just a Google registration.

Opportunities

In the time I have remaining, let me look at two different doors that are now opening. Each door is a metaphor for a way to exploit Google as a game platform as well as a platform for building game like applications. In short, I want to suggest that the notion of a game must be viewed in two ways.

First, I think it is wise to look at the Google platform as one in which 10 percent of one's development effort should be invested. The reason is that Android and Wave are immature or not yet built out. Therefore, the idea is to take a simple game idea or an existing game feature and recycle it for Google. Within the next two to five years, additional development resources should be directed at the Google platform. Five or six years out, development for only the Google platform is likely to be possible.

The reason for this is that the game platform and game device market does not change as rapidly as some believe. The high end, dedicated game devices will persist in the market. In the short term, the Apple iPhone is a more viable mobile game platform. However, over time, the shift to mobile computing and cloud computing will change the equation. In short, learn and recycle. Don't bet the farm on Google.

Second, I think it is important to recognize that Google moves in small, incremental steps. The company does this in order to avoid alerting competitors to its broader strategy and to minimize antitrust actions. Nevertheless, you should plan on allocating your time based on how the Google market shapes up. This means that delay in learning how to code for Google is a bad idea. Among the technologies to learn are SketchUp (Google's drawing program), Android (the visible part of the Google operating system), Wave (collaborative spaces), and Google Apps and OneBox APIs. These functions are, at a minimum, the way in which to obtain the Googley expertise you need.

In closing, let me make three observations about Google, games, and the game like applications that will be the norm in computing in the future:

First, Google is a semi open source play. As a result, Google has a viable chance to choice out proprietary approaches to certain types of software. Apple and Microsoft are closed. Open may become a more significant factor and quickly, so be prepared.

Second, the game sector itself is changing. The increase in mobile and cloud computing will open new doors. Google is a cloud company; therefore, learning how to make Google do your bidding is an investment in where next generation computing is moving.

Finally, the game approach will find its way into automatic teller machines, automobile dashboards, robots and drone war fare, tools for medical procedures, and other business sectors.

Betting on Google is a good idea. Because computing is becoming game like, you want to build for the platform that will command a significant market share. Google is a must learn platform, and it is squarely in the game mainstream.

Author's Note: This is an expanded outline of my remarks. I have not included the PowerPoint slides, tables, and other data in this summary. These will be made available on the Beyond Search Web log on the day of the talk in July 2009.

.

GAMING CONCEPTS IN ACCESSIBLE HCI FOR BARE-HAND COMPUTER INTERACTION

Kristopher Reese, Ahmed Salem, and George Dimitoglou
Department of Computer Science
Hood College
Frederick, MD, USA
{kwr2, salem, dimitoglou}@hood.edu

KEYWORDS
Accessibility, Human-Computer Interaction, Human Factors, Computer Vision, Bare-Hand Controls, Object Detection, Future Interfaces

ABSTRACT

Hand-Computer interaction is a frequently researched topic in the field of computer gaming. However, very few of the devices or techniques used for Hand interaction with are accessible within many smaller groups, especially those with disabilities. For this paper, we have developed a tool to show several of the techniques used in gaming research and describe how these techniques could be implemented in Human-Computing Interaction with a focus on accessibility for the physically challenged. A Bare-Hand tracking technique is used to track the location of the Hand in relation to a single web camera. This proof of concept interface offers a look at how future devices could be both fun for typical users without sacrificing on accessibility.

1. INTRODUCTION

The recent decade has brought growth and innovation to the computer gaming industry. Much of the innovation has been focused on human-computer interaction and techniques to enhance how users interface with video games. New interfaces, such as the Nintendo Wii-mote and the Sony Eye-Toy, are the results of more than 20 years of research in the fields of Virtual Reality and Spatial 3D Interactions (LaViola 2008). These technologies now dominate gaming but their potential applications to other areas have hardly been met. In particular, applying these technologies to provide enhanced computing accessibility to those with physical disabilities. For example, replacing computer mice on computer systems or implementing virtual input devices replacing keyboards for those with motor skill disabilities that are unable to use physical input devices.

Such technologies often remain inaccessible to those with some forms of disabilities. In 2004, the United Nations released a report with statistics about disabilities at an international scale, stating that in most countries at least one in ten people are disabled by physical, mental, or sensory impairment and that at least 25 percent of any population is adversely affected by the presence of disabilities (United Nations, 2004). Population aging also augments these numbers by introducing computer users with degraded body functions, including vision, memory, and motor skills. According to the World Health Organization, there were 600 million people above the age of 60 in 2000, and it is expected that the number will double by 2025 (World Health Organization, 2004).

It is clear that new and innovative forms of interfacing with computers are becoming increasingly necessary both to those with disabilities and to an aging but computer savvy population. While current technologies have a lot of potential providing a wide variety of solutions, we focus on techniques that allow users to interact with computers strictly by motion.

The following section provides an overview of related work in the areas of assistive technologies for game interfaces and later covers the background work specific to object and gesture detection. The next section describes a new method for hand detection that allows human-computer interactions using a web camera. Subsequently, the architecture of the Motion Control Application (Mocoa) is described, along with the algorithm that detects a user's hand and how the overall system operates. Recognizing the importance of performance and how it may affect the effectiveness of an interfacing application a brief performances analysis section illustrates key areas that can enhance the utilization of the application. The last section includes the conclusion and discussed future work.

2. RELATED WORK

As many of these devices begin their development in the gaming industry, looking at the development and accessibility becomes a necessity. By looking at Gaming Accessibility we can observe the development of these devices beyond the typical user interactions that they were developed. However, Grammenos et al. make clear observations that relatively few efforts have been devoted to game accessibility and those with the primary concern of accessibility in gaming are organized groups of disabled people (Grammenos et al 2009).

Grammenos et al. later states that there are often two main approaches to addressing the issue of game accessibility, which often applies to fields outside of gaming as well. Grammenos et al. states that inaccessible games are often made accessible through third-party assistive technologies or are developed from scratch, often with a specific disability in mind (Grammenos et al. 2009). There are drawbacks to both approaches; however, these drawback would be limited if new forms of input devices are developed that would be easy to use and developed for users with both disabilities and for those without disabilities. The Nintendo Wii-mote and the Sony Eye-toy are a step in the right direction, but more work needs to be done to make them more usable for a larger percentage of the disabled population.

LaViola's article explores the history of Video Game interaction and discusses some of the future developments of spatial 3D Interactions. LaViola recognizes that though these devices work well enough for current gaming, future developments in these technologies must occur in order to keep up with computing and become useful in fields outside of gaming. (LaViola 2008) There are limitations in the current technologies that prevent the use of these technologies outside of gaming. The Sony Eye-Toy, for example, was built to detect hand gestures and does not measure a third dimensional space. We see similar limitations in the Nintendo Wii-mote, which will detect six degrees of freedom with conventional methodologies and handle only exaggerated gestures reliably.

Unlike commercial devices, the web camera offers a cheaper and often just as effective device for hand-computer interactions in gaming. The web camera is also a common device used in the research for bare-hand computer interaction. Because of this, we must look at research related to bare-hand interaction in relation to web cameras.

One often employed method of using the webcam as an input interface is through the use of estimation of motion vectors. This technique is proposed through the use of a block matching method. (Jain and Jain, 1981; Sun and Cheng, 2007) We also see similar techniques used in a three-step search (Li et al., 1994), four-step search (Po and Ma 1996), and a diamond search (Zhu and Ma, 1999). These techniques offer an effective method for tracking objects that move in front of the web camera, but have proven ineffective in many other situations. In these methods, a simple change in lighting or an object moving into the plane may cause miscalculations and ultimately lead to false positive motion tracking.

Another frequently used method for detecting hands and gestures is through the use of a form of glove with varying colors, a method introduced by Frederickson (Frederickson et. al, 2008). This technique creates hotspots on the hand by giving each finger a specified color. It also adds a marker to allow the computer to determine the x-, y-, and z- coordinates. This practice allows a user to use many different hand positions and gestures. The glove is also low cost as there is no special hardware inside of each glove. However, the use of this technique does not allow a user who is not able to put on a glove the use of this technique. For this reason, a bare-hand control interface would be best suited to allow accessible usage for the disabled to computing and gaming

There has been significant progress in bare-hand control interface research in the last few years. Lee and Lee as well as Hardenberg and Bérard have built bare-hand detection algorithms that can be used for tracking the motion of a users hand. Lee and Lee have built an algorithm that detects the hand and head using a color difference technique (Lee and Lee, 2008a; Lee and Lee 2008b). For this technique to work, Lee and Lee have assumed human skin color is always of a red tone. This, however, is not always true since human skin colors vary in tone and hues. This assumption will work for the majority of skin types but not with all skin types.

Hardenberg and Bérard propose a different method for the detection of a hand. They developed an algorithm to detect user fingers using the basic shape of the tip of a finger (Hardenberg and Bérard, 2001). They then use a motion vector tracking algorithm to determine which objects, in this case fingers, have moved and to what locations. This algorithm, however, is dependent on a user's ability to move fingers. Because of this constraint, the algorithm is not suited for all cases of accessibility, especially for those physically disabled with quadriplegia.

Due to lack of a full interface that put accessibility above other goals, it became necessary to create a new framework that met specific requirements for accessibility purposes. This framework would need to meet several basic requirements that would allow a physically disabled user the ability to use this interface. However, meeting these requirements would not make the interface usable by users. For this, several advanced requirements would need to be addressed in order to

3. FRAMEWORK ARCHITECTURE

The Mocoa (Motion Control Application) project is developed as an accessibility framework to enable the integration of user gesture and motion accessible applications. Mocoa is structured so that any application could be easily integrated into the system as a software module. Figure 1 shows the top-level architecture of the project.

The software contains three primary components, the web camera controller, the object detector component, and the user interface. The web camera controller takes

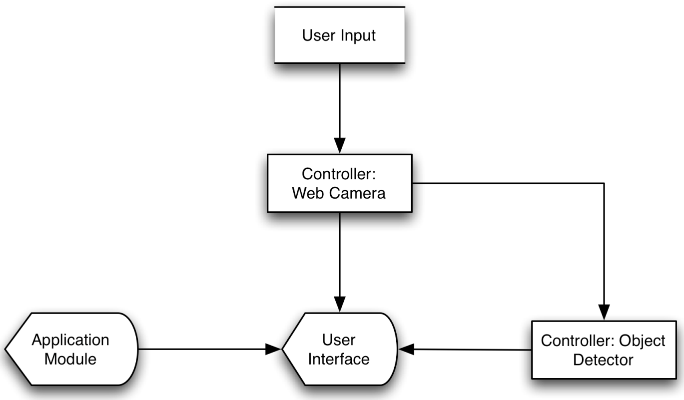


Figure 1: Basic Structure of Mocoa

in a captured frame as user input. This captured frame is then sent to both the User Interface, for display to the user, and the object detector component where the primary detection of the hand is completed. The User Interface then receives the information from the Object Detector. This information is sent between an application module and the User Interface. In order to ensure that the system is usable and accessible to users of all physical capabilities, the interface must meet several basic requirements, as well as suitably meet several advanced framework requirements.

3.1. Basic Framework Requirements

The basic framework requirements serve as the minimal requirements for making a system, which offers bare-hand interaction with a computer. Hardenberg and Bérard identify several different restraints for a vision-based human computer interaction system. These requirements consist of: Detection, Identification, and Tracking (Hardenberg and Bérard, 2001). Mocoa offers a unique method for object Detection and Tracking using a web camera as the input device.

3.1.1 Detection

Detection is necessary to determine the presence of an object within the frame of the captured image. Mocoa uses the Viola-Jones algorithm (Viola and Jones, 2001) for detecting the presence of various objects. This object detection algorithm is flexible enough to enable changes between control objects such as a head or a fist. This flexibility that requires only a brief calibration exercise offers great versatility and allows use by a diverse pool of users with different types of disabilities,

3.1.2 Identification

Object identification is an important factor for a hand motion controller. Without some form of identification, determining which object is the controlling object becomes difficult. This difficulty becomes more obvious as more objects are introduced into the field of view. There are other types of identification that are necessary for determining an object. In Mocoa, the following types are necessary:

- Hand posture Identification. Mocoa uses the hand posture of a sign language letter "A" to identify as an object within the frame. However, other object identifiers, such as a user's head or various other body parts, can serve as an identifier.
- The 2-dimensional location of the object. Mocoa uses this information to move a cursor or other virtual object based on the *x*- and *y*-coordinates of the real object.
- A third dimensional space. This offers a more Realistic experience with the computer by allowing the user to interact with a virtual object as they would with a realistic object.

Mocoa achieves identification through the use of the classifiers to differentiate between various objects in the image. Since a user's hand is a 3-dimensional object, for this project we have limited the identification of the object to a sign-language letter "A" gesture. After this object is identified, it returns the central location of the detected object as *x*- and *y*- coordinates.

3.1.3. Tracking

Once the object is detected and identified, it has to be tracked to allow the "registering" of the user's motion as a controlling motion in order to control various aspects of computing. Mocoa achieves a form of tracking by restarting the object detection and identification process during each captured frame. This helps to ensure that the stability of the input remains consistent throughout the remainder of the user's input. This may have caused some issues related to performance, but the benefits of tracking an object in this way far exceeded the loss of performance.

3.2. Advanced Framework Requirements

The basic framework requirements ensure that the framework is able to accommodate basic vision-based human computer functionality. However, to make an application useable in a practical way, there are advanced requirements that need to be met. These advanced requirements consist of input stability, spatial resolution, and input latency.

3.2.1. Input Stability

Input instability manifests when ambient lighting, motion of other objects, or electrical noise become present in the user's environment and interferes with Mocoa's components. The result of any of these problems may cause erratic behavior of the controlled component (e.g. mouse pointer) or false positives in hand detection. If the input mechanism does not remain stable, oscillations could occur making the interface difficult to control. However, even if they occur, Mocoa is resilient enough to allow the instability to last only a few frames, except in extreme instances where an object may have a similar shape as the object or where lighting is not great enough to detect the hand.

3.2.2. Spatial Resolution

Spatial Resolution becomes necessary when the resolution of the captured frame is much smaller than resolution of the device that is being controlled. Hardenberg and Bérard suggest that for a point-and-click system, the smallest possible pointer movement must be no larger than the smallest selectable object on the screen. Mocoa handles spatial resolution through the use of its calibration tool. The calibration tool allows a user access to set up a maximum and minimum coordinate for movement. This information then translates movements into percentages, which can then used to control an application with variable pixel sizes.

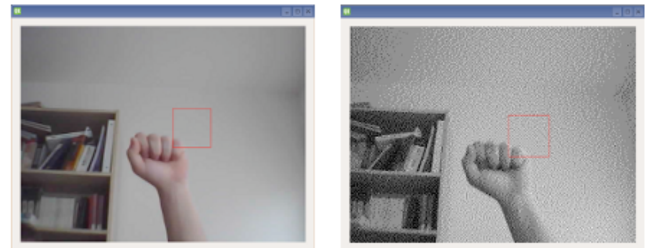
3.2.3. Input Latency

Input Latency is one of the most significant advanced framework requirements that have an impact on overall usability of the interface. High latency times from the input to the running module would significantly degrade the user's experience of the interface. For any form of input device, especially in gaming, low latency inputs offer players the ability to control the system with a more natural feel. Mocoa achieves low latency times through the use of the detection algorithm. The algorithm used achieves a high rate of success in a short amount of time. This, coupled with the stability of the tracking system, creates a low latency interface, which creates a very natural feel when controlling virtual objects.

4. HAND DETECTION

The algorithm used by Mocoa for hand detection and identification is the Viola-Jones algorithm (Viola and Jones, 2001). The Viola and Jones algorithm was designed primarily for face detection, but has proven to be just as effective in a search for other objects when the haar classifier files are changed. Unlike other algorithms (Lee and Lee, 2008; Lee and Lee, 2008b), the use of an object detection algorithm eliminates the reliance on the color or hue of the skin.

The Viola-Jones algorithm is an implementation of the AdaBoost machine-learning algorithm. The Viola Jones Algorithm works by taking areas of the image and scanning the image for various sum calculations based on an integral-image cascade classifier file. In our implementation, these classifier files are an XML based



(a) *(b)*

Figure 2: Hand Detection in both a color image (a) and a black and white monochrome image (b)

file that allows the Viola-Jones algorithm to determine if the object we are searching for is located within the portion of the image that is being scanned. This is repeated for all portions of the image.

The implementation of the Viola-Jones used by Mocoa does not force users into a specific hand gesture or require users to be physically able to use a specific body part by allowing a user to easily change the haar classifier file. Through the use of the Viola-Jones algorithm we can now also reliably detect an object under various lighting instances without the requirement of a color image. Figure 2 shows that by using the Viola-Jones algorithm, Mocoa is able to detect a hand in both a color and a black and white frame. The algorithm, also addresses several other issues that had been identified by Lee and Lee such as the issue of overlapping body parts; for example, when a users hand overlaps with their face in the Lee and Lee algorithm, the face is fallaciously identified as the controlling object along with the users hand

One of the primary issues with Lee and Lee's detection algorithm is that overlapping body parts (with the same skin color) will prevent a hand from being detected. The Viola-Jones algorithm implemented in Mocoa allows for the collision of various body parts without degrading the recognition of the object in the frame. This is achieved by using the Viola-Jones detection algorithm and the lack of need for identification using skin color. Figure 3 shows the detection algorithm when the hand and a user's face are close or overlapping.

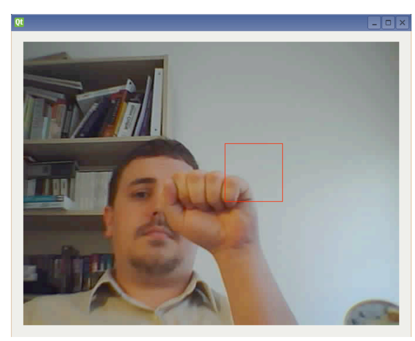
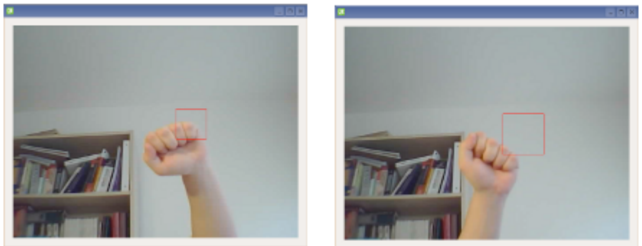


Figure 3: Overlapping Face and Hand, an issue seen in (Lee and Lee 2008a; Lee and Lee, 2008b), is resolved.



(a) *(b)*

Figure 4: Hand Detection with 45 degree turn to the left (a) and to right (b).

The Viola-Jones algorithm is able to quickly identify objects in the frame; however, simply using the Viola-Jones algorithm will not detect objects that have been rotated. Lienhardt and Maydt extended the Viola-Jones algorithm to allow the algorithm to detect objects that have been rotated up to 45 degrees from standard orientation (Lienhardt and Maydt, 2002). The Viola-Jones algorithm implemented by Mocoa uses the extension to the algorithm made by Lienhardt and Maydt. Figure 4 shows the object detected at approximately 45 degrees in both directions. By implementing this extension, we allow users the ability to naturally turn their hand as they move without causing loss of object detection.

The detection algorithm is able to determine the 2-dimensional location of the hand. Mocoa uses the central location of the hand to control aspects of computer. The detection algorithm used has been fitted to calculate the *x* and *y* coordinates of the bottom left corner of the box that wraps the hand when detection is successful as well as the height and width of the object. In using this point, as well as the height and the width of the box, we can calculate the *x*- and *y*- coordinates for the center of the box. Equations 1a and 1b show the central *x* and *y* coordinate.

$$f(x) = \frac{rect_x + rect_w}{2} \qquad f(y) = \frac{rect_y + rect_h}{2}$$

(a) *(b)*

Equation 1: Mathematical equation for calculating the central x (a) and central y (b) coordinates which are used for moving a virtual object.

Calculating a third dimensional space, which could be beneficial inside many applications, becomes a bigger challenge. However, simple mathematics can be used to determine a *z*-coordinate based on the size of the object detected with relation to a calibrated 0 point. Equation 2 is a mathematical representation of the algorithm used to calculate the *z*- coordinate in Mocoa, where h_c is the maximum calibrated height in Mocoa and h_0 is the current height of the detected object within the frame. "*sf*" signifies that the results of the equation should be floored to the number of significant figures given in the equation.

$$f(z) = \begin{cases} +z & if \left\lfloor \frac{h_c}{h_0} \right\rfloor_{sf=1} > 1 \\ 0 & if \left\lfloor \frac{h_c}{h_0} \right\rfloor_{sf=1} = 1 \\ -z*10 & if \left\lfloor \frac{h_c}{h_0} \right\rfloor_{sf=1} < 1 \end{cases}$$

$$where \quad z = \left\lfloor \frac{h_c}{h_o} \right\rfloor_{sf=1}$$

Equation 2: Mathematical representation of the z-coordinate algorithm used for calculating 3-dimensional space on the 2-dimensional framed based on relative sizes.

The information received through the calculations in three-dimensional spaces is then transmitted to requesting modules. However, the resolution of a web camera may be very small in comparison to the resolution of a module; for example, a web camera may only capture a 600 pixel by 600 pixel frame while the resolution of the operating system is likely to greatly exceed the frame resolution of the web camera, potentially 1600 pixel by 1200 pixel, or greater. Because of this, the coordinate information received from the object detector component is checked against the calibration tool to calculate a percentage value. This is then used to move a virtual object with relation to the maximum and minimum *x*- and *y*- coordinates.

5. PERFORMANCE ANALYSIS

Performance of an interface plays a major role in the usability and accessibility of an application. Poor performance may lead to higher latency times and a degradation of the overall stability of the system. Mocoa was initially designed as a single-threaded application. During development, Mocoa began to show signs of high computational needs, which ultimately caused issues with latency times and the overall usability of the interface. Because of this, Mocoa was later made multi-threaded to take advantage of the multi-core processors that are becoming more prevalent in home computers today. This performance analysis was conducted to determine the most suitable environment for this interface to exist.

To conduct this analysis, it was necessary to integrate a module that could be controlled using the Mocoa framework. For this, a simple version of one player Pong was integrated into the system. This module took advantage of each of the components inside of Mocoa so that the entire system could be analyzed for performance related reasons.

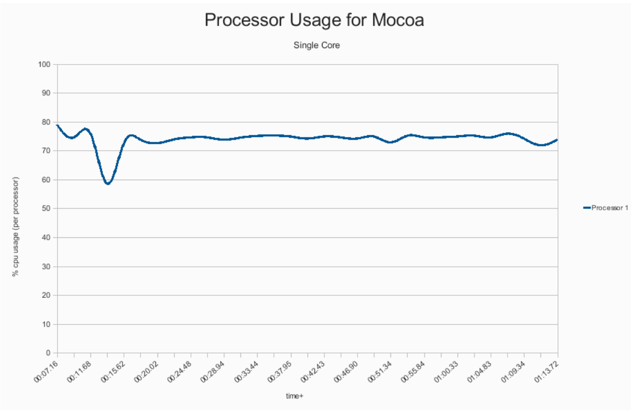


Figure 7: Single Core Processor Usage Performance Analysis

This performance analysis was conducted on a Linux machine running a standard installation of Ubuntu 8.10 64-bit version using a "Creative Live! Notebook Pro" web camera with the V4L2 drivers for Linux webcams. We run the analysis on a single-, dual-, and quad- core processor running at the same clock speed and cache per processor. Table 1 shows a more detailed description of the systems used for each of the system performance analyses.

	Computer 1	Computer 2	Computer 3
OS	Ubuntu 8.10	Ubuntu 8.10	Ubuntu 8.10
Clock Speed	2.66 GHz	2.66 GHz	2.66 GHz
Cache	2 MB	4 MB	8 MB
FSB	1066 MHz	1066 MHz	1066 MHz
Cores	1	2	4
RAM	4GB	4GB	4GB

Table 1: The hardware specifications used for the performance analysis of Mocoa

We began this analysis by measuring the average CPU usage over a minute of typical usage of the operating system. In all cases, the operating system ran at approximately 20% of the processing with a standard deviation of 1-2%. Knowing the processor overhead, we can begin to analyze all of the information received from our processing dumps for each core of the processor. For the analysis, the pong module was played for 1 minute while the processor usage of the application was analyzed.

Our first analysis was conducted on a single core system. Figure 5 shows a graphical representation of the results of this analysis on a single core processor. The user saw significant latency issues with both the game and user input as the processor attempted to calculate the user's input on a processor running at nearly its full capacity. We see that the average processing power necessary on a single core processor is approximately 74%. With the average Operating System overhead, we can conclude that the latency issues within the application were likely due to the processor running at nearly 100%.

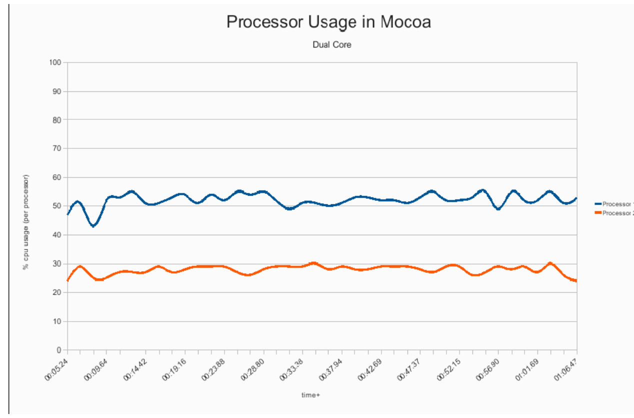


Figure 5: Dual Core Processor Usage Performance Analysis

After Multi-threading the interface, we see significant improvements in the latency times between hand movement and virtual interaction. While running on multiple cores, we see significant improvements in the processing needs for the application. Figure 6 graphs the processor usage of Mocoa on a dual core processor. We can see that on a dual core processor, Mocoa uses just over 50% of the first processor while secondary necessities require 27% of the second processor. This results in a much lower latency time, allowing the user to more naturally play a game or control other aspects of computing.

The quad core analysis shows insignificant changes in the processing needs of the application. The first core remains unchanged from the dual Core analysis, running at just over 50% of the processing needs. Each subsequent core ran at 19% of the processor cores. No latency changes were noticed between runs on a quad core and a dual core.

Through these analyses, we can assert that it is becoming more feasible to implement a Bare-Hand computer input interface, such as Mocoa, which requires high processing power, in many aspects of computing, including home computing and gaming.

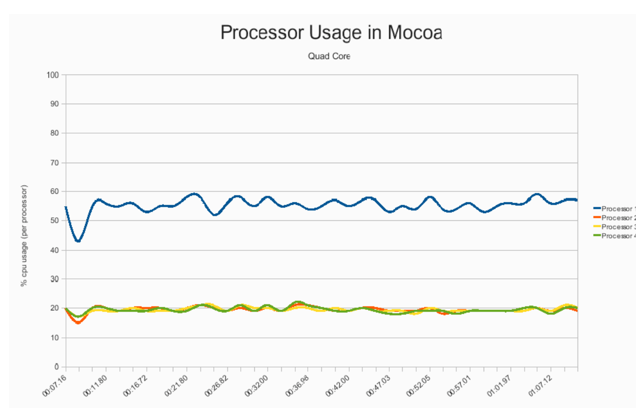


Figure 6: Quad Core Processor Usage Performance Analysis

6. CONCLUSION & FUTURE WORK

Mocoa serves, as a tool that shows that computing today is capable for handling an interface that can be used by a variety of users, including those with physical disabilities. One idea for future development of Mocoa is to integrate this technology in specialty devices that are used by patients with certain disabilities such as quadriplegia.

In certain forms of quadriplegia, the disabled have limited control over the movement of their hand, but little to no control over the movement of fingers or the strength at which they hit the various input devices. Because of this, keyboards and other input devices that are frequently touched need to be replaced often.

Mocoa offers a unique solution for hand-computer interaction controls. It resolves many of the issues that are found in various other algorithms and successfully addresses issues that were present with different systems related to user characteristics such as skin color or saturation. Still, Mocoa has its own flaw: the problem of false positives, due to surrounding image noise, which remains unresolved. However, we are optimistic that existing techniques can be used to reduce the noise and address this current shortcoming.

With the success of net-book computers and with mobile devices, such as cell phones, becoming more able to handle complex computing, it would also be beneficial to determine where improvements in the implementation of Mocoa and its object detection algorithm so that this interface might be used with these types of low resource devices.

Overall, Mocoa shows promise in the deployment of hand-based computer interaction, but more work remains to be done to take what currently is a useable proof-of-concept to a more comprehensive and feature-full framework.

ACKNOWLEDGEMENTS

We are grateful to Mr. James Devilbiss for several of his ideas with respect to accessible computing applications for this project.

REFERENCES

Fredriksson, J., Ryen, S. B., and Fjeld, M. 2008. Real-time 3D hand-computer interaction: optimization and complexity reduction. In Proceedings of the 5th Nordic Conference on Human-Computer interaction: Building Bridges (Lund, Sweden, October 20 - 22, 2008). NordiCHI '08, vol. 358. ACM, New York, NY, 133-141.

Grammenos, D., Savidis, A., and Stephanidis, C. 2009. “Designing universally accessible games.” Comput. Entertain. 7, 1 (Feb. 2009), 1-29.

Hardenberg, C. von and Bérard, F. 2001. “Bare-hand human-computer interaction.” In Proceedings of the 2001 Workshop on Perceptive User interfaces (Orlando, Florida, November 15 - 16, 2001). PUI '01, vol. 15. ACM, New York, NY, 1-8.

Jain, J.; Jain, A.; 1981. “Displacement Measurement and Its Application in Interframe Image Coding” In Transactions on Communications, IEEE (Dec. 1981): 1799-1808

LaViola, J.J., 2008. “Bringing VR and Spatial 3D Interaction to the Masses through Video Games.” Computer Graphics and Applications, IEEE. Vol.28, no. 5 (Sept.-Oct.): 10-15.

Lee, Dae Ho and Young Jae Lee; 2008a. “Sensing and Motion Control of Virtual Objects for Web Camera-Based Game.” In Proceedings of the Second Conference on Future Generation Communication and Networking. (Sanya, China, Dec. 13-15) FGCN: 28-33.

Lee, Young Jae and Dae Ho Lee, 2008b. “Research on Detecting Face and Hands for Motion-Based Game Using Web Camera.” In Proceedings of the 2008 International Conference on Security Technology. (Sanya, China, Dec. 13-15). SECTECH: 7-12.

Li, R, B. Zeng, and M.L. Liou, 1994. “A new three-step search algorithm for block motion estimation.” In Transactions on Circuits and Systems for Video Technology, IEEE (4), 438-442.

Lienhart, R.; Maydt, J., 2002 “An extended set of Haar-like features for rapid object detection.” In Proceedings of the 2002 International Conference on Image Processing. (Rochester, NY, Sept. 22-25). IEEE: I-900-I903.

Po, L.M; Ma, W.C. “A novel four-step search algorithm for fast block motion estimation” In Transactions on Circuits and Systems for Video Technology, IEEE (June 1996): 313-317.

Sun, Huey-Min and Wen-Lin Cheng; 2007. “The Input-Interface of Webcam Applied in the Digital Content Interaction” In Proceedings of the Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing. (Splendor Kaohsiung, Taiwan, Nov. 26-28). IIH-MSP: 2007.

United Nations, 2004. World programme of action concerning disabled persons: Current Situation. <http://www.un.org/esa/socdev/enable/diswpa04.htm>

Viola, P.; Jones, M., 2001. “Rapid object detection using a boosted cascade of simple features” In Proceedings of the 2001 Conference on Computer Vision and Pattern Recognition. (Kauai, Hawaii, Dec. 9-14). CVPR: I-511-I-518.

World Health Organization, 2004. Department of Chronic Diseases and Health Promotion (CHP): About aging and life course. <http://www.who.int/ageing/en/>

Zhu, Shan; Ma K.K., 2000 “A new diamond search algorithm for fast block-matching motion estimation.” In Transactions on Image Processing, IEEE (Feb 2000): 287-290

TOWARD A CAUSAL MODEL FOR AUTOMATIC GAME BALANCING

Jeffrey Moffett, Charles Rich and Joseph Beck
Worcester Polytechnic Institute
100 Institute Road
Worcester, Massachusetts, USA
{jeffmoffett | rich | josephbeck}@wpi.edu

KEYWORDS

Causal model, game balancing, fun, AI, machine learning, optimization.

ABSTRACT

We have started to develop a mathematical model of how various aspects of a computer game influence how much fun it is to play. This causal model is organized into three layers: a generic layer that applies to any game, a refinement layer for a particular game genre, and an instantiation layer for a specific game. We have conducted an initial user study to validate the model using a simple sports game called Slime Volleyball. In addition to providing insight into game design, our causal model is intended to provide a foundation for the application of machine learning and optimization techniques to improve the player's experience.

INTRODUCTION

Recent efforts toward applying machine learning and optimization techniques to computer games, such as first-person shooters (Hefny et al. 2008) and real-time strategy games (Ulam et al. 2008), have assumed that the goal is to achieve the best possible performance of the AI-controlled nonplayer characters (NPCs). In fact, game design books (Adams and Rollings 2007) warn against exactly this fallacy. The goal of applying AI to games *should* be to improve the player's experience—an overly dominant opponent is not a good experience!

Currently, what is broadly referred to as "game balancing" is done manually during the play testing phase of game development. In manual game balancing, the many parameters of a game are adjusted by trial and error to test their effect on player experience, without an overall guiding model of the dependencies between parameters. This approach is both inefficient and cannot support automatic balancing during game play, for example, to adapt to particular players as they change their skill levels.

Viewing automatic game balancing as an optimization problem, we must have an explicit objective function with respect to which we can optimize. To use a simple word, we call this objective function "fun". Our causal model of fun will be useful both to support automatic game balancing and to provide insight into how different game design decisions are likely to affect the player experience.

To validate our proposed model, we conducted a pilot study in which 136 people played a simple sports game, called Slime Volleyball, under different parameter configurations and reported on their experience.

In the future, we plan to apply automatic optimization techniques to our causal model to automatically adjust the parameters of Slime Volleyball and other games in order to keep the player in his/her "fun zone."

RELATED WORK

Ours is not the first attempt to develop a formal model of computer games. However, compared to previous work, our model is both more detailed and more practical for application to automatic balancing of games.

Explaining Fun

The EVE model (Burns 2007) explains fun in terms of Bayesian information theory. It is comprised of three components: Expectation, Violation, and Explanation. EVE argues that a positive experience is produced when either a situation matches the player's expectations or a situation violates the player's expectations, but s/he has an explanation for why the violation occurred. EVE has been applied to the analysis of gambling games (Burns 2006), music creation (Burns 2006a) and computer games (Piselli 2006).

Modeling Interaction

MDA (Hunicke 2005) was designed as a way to understand the interactions involved in a computer game. A game is decomposed into three components: Mechanics, Dynamics, and Aesthetics. The mechanics of the game are the low-level data structures and algorithms. Dynamics are the run-time interactions between the user and the system. Aesthetics are the emotional responses of the player while playing the game. MDA has been used to create a game-monitoring system, called Hamlet, which attempts to improve the player's aesthetic experience in combat-based computer games. The system dynamically estimates the player's probability of dying based on the previous enemy encounters and intervenes to make the game easier when needed. Initial results showed that Hamlet decreased the number of player deaths while keeping its operation virtually undetected by the player.

Qualitative Flow

The GameFlow model (Sweetser 2007) provides a qualitative approach to evaluating player enjoyment of computer games, using the well-known concept of "flow" (Csíkszentmihályi 1998) as the starting point. Each of the eight elements of flow are mapped to typical game elements. GameFlow also adds a ninth element, social interactions, which exists in some games but is not included in the standard flow model. The GameFlow model has been used to analyze player enjoyment in two real time strategy games, Warcraft 3 and Lords of Everquest.

Heuristic Entertainment Value

Yannakakis and Hallam (2007) developed a heuristic function for entertainment value in predator/prey games, such as Pac-Man. By extensively analyzing this type of game, three elements were theorized to directly effect entertainment value: challenge level, predator behavior diversity, and predator spatial diversity. The heuristic function was validated by comparing user surveys of game sessions with the values predicted by the heuristic function. Of the related work mentioned above, this is the most similar to ours.

A CAUSAL MODEL

A causal model (Cohen 1995) is a directed acyclic graph in which nodes represent model factors and each edge indicates that the source factor directly influences the target factor. For example, our causal model in Figure 1, which we will explain in more detail below, has the node labeled *fun* at the root of the graph. According to the model, the factors labeled *player*, *ambience*, *game metrics* and *perceived fairness* directly influence fun. Other factors indirectly influence fun through their effect on these factors. A causal model is useful since it suggests experiments to conduct to test how various factors affect others.

Our causal model contains three layers: a generic layer that applies to any game, a refinement layer for a particular game genre, such as sports games, and an instantiation layer for a specific game, such as Slime Volleyball.

Generic Layer

Even though the root node of the generic layer of our causal model is labeled *fun,* we are not trying here to make any deep claims about the philosophical or psychological nature of the human experience of fun. In the pilot study described below, the degree of fun will simply be determined by asking the player a rating question. Fun could in principle be determined in other ways, such as by observation of the player by judges, or measurement of biological signals.

Our generic layer covers both games in which the player plays only versus the environment and games in which there is an NPC opponent. The specific game we chose for initial validation of the model involves an NPC. For a game without an NPC opponent, all the NPC-related nodes in the model can simply be ignored.

The first factor directly influencing fun is the *player.* Players vary in a great many ways, including, age, gender, and gaming experience. In applications of this model, the player is not usually a variable that can be manipulated.

The second factor directly influencing fun is the game *ambience,* in which we include all of the artistic, stylistic and aesthetic elements of the game, such as graphics and music. Even though these aspects of a game undoubtedly affect how much fun it is to play, this is not an area we have yet explored. Thus there are currently no nodes in the model that directly influence ambience.

The third factor directly influencing fun is the *game metrics.* Game metrics include all of quantities that the game code keeps track of, such as score, duration, kill ratio, accuracy and so on.

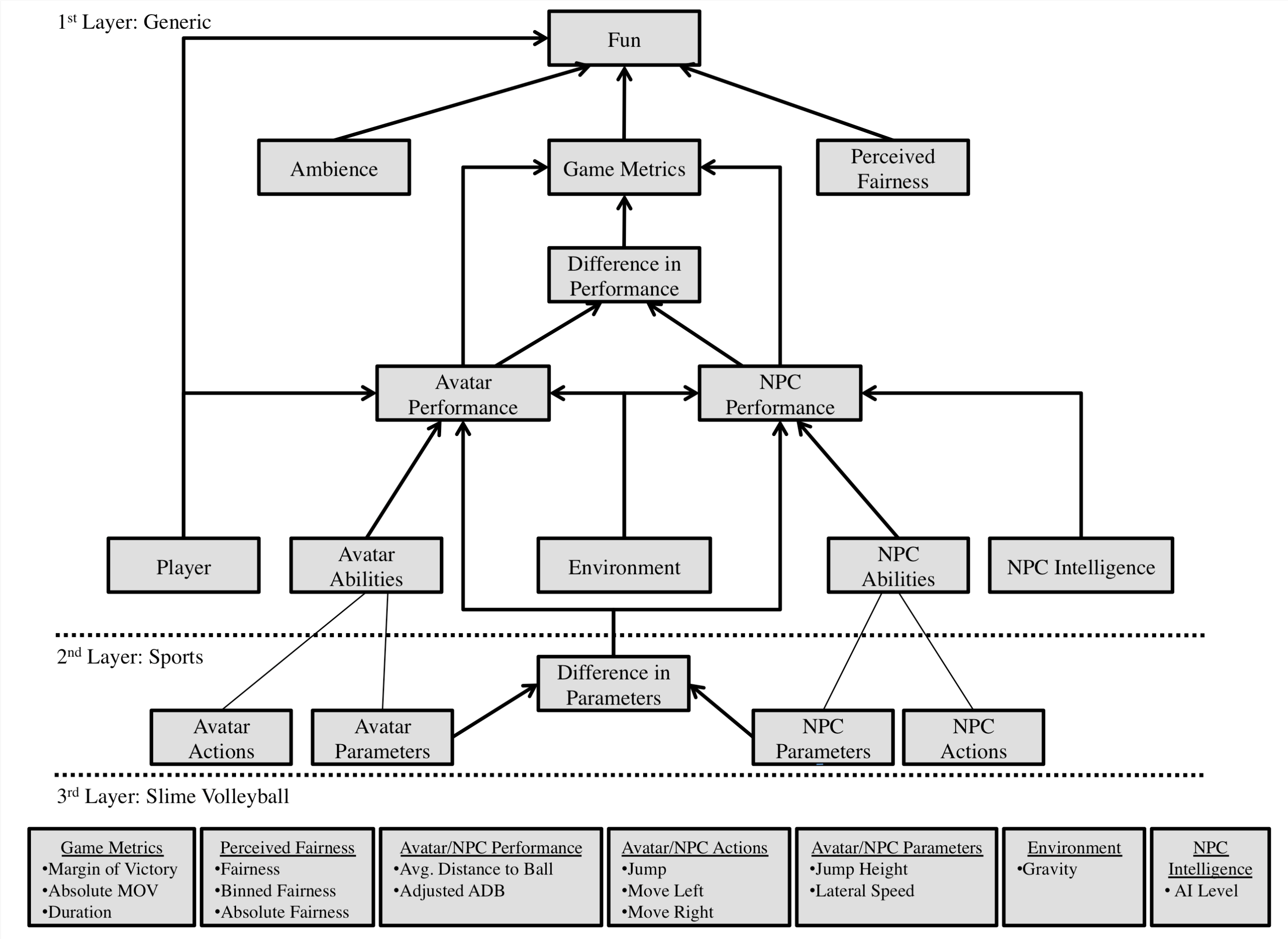


Figure 1. Three Layer Causal Model

The fourth factor directly influencing fun is *perceived fairness*. Players want to believe that they not only have a chance of winning, but also a chance of losing. It's not fun to play a game in which either you or your opponents have an unfair advantage. In our initial construction of the generic layer, we were unsure of the direct influences on perceived fairness. In the pilot study, however, we discovered a strong influence from game metrics to perceived fairness in Slime Volleyball (SVB).

Since game metrics are used to judge performance, two of the direct influences on game metrics are the *avatar* (character controlled by the player) *performance* and the *NPC performance*. Also, in many games, what matters most is not the individual performances of the avatar and NPC, but their relative performance. We added the *difference in performance* node to the model to reflect this influence.

Since the player controls his/her avatar, there is an obvious direct influence from the player to avatar performance. The performance of the avatar is also typically influenced by a set of *avatar abilities*, the details of which depend on the game genre and specifics.

Symmetrically, the NPC performance is directly influenced by the *NPC abilities.* Corresponding to the player's control of the avatar, the *NPC intelligence* controls and therefore directly influences the NPC performance.

Finally, the *environment* node represents everything in the game world that is not an avatar or NPC. Examples of environmental properties include are terrain, gravity, spawning locations and weather. In general, changes in the environment directly influence both player and NPC performance.

Genre Layer

The second layer of abstraction in our causal model refines the generic layer with concepts and distinctions related to a particular game genre, such as first-person shooters, real-time strategy games, or sports games. The refinement process may include both adding additional nodes and edges and expanding/splitting a single node into one or more nodes.

Figure 1 shows the refinement of the generic layer for the sports game genre. For sports games, the avatar and NPC abilities nodes are each split into *actions* and *parameters*. Actions are the fundamental vocabulary of interactions a character can have with the game world. Parameters constrain these actions. For example, jumping is an action; how high you can jump is a parameter.

In many sports games, the avatar and NPC have the same actions, so that what is important is the difference in their parameters. We have therefore added a *difference in parameters* node with direct influences to both avatar and NPC performance.

Game-Specific Layer

The relationship between the genre layer and the game-specific layer is different than the relationship between the generic first layer and the genre layer. In the game-specific layer, we instantiate the abstract concepts in the generic or genre layers with one or more concrete quantities or factors in a specific game. Figure 1 illustrates the game-specific layer for SVB.

SVB is a two-dimensional volleyball game (see Figure 2) in which the player controls the creature (slime) on one side of the net and an AI controls the other. Each creature can move laterally and jump to block and return the ball. As in volleyball, you lose the point if the ball touches the ground on your side of the net. A match ends when either side scores five points.

Starting on the left of the third layer, we see three concrete game metrics candidates: *margin of victory* (avatar score minus NPC score), *absolute margin of victory (*absolute value of margin) and *duration* (how long the player played the game before quitting).

For perceived fairness, there are also three concrete candidates: the raw *fairness* value (0-9) provided by the player, the *binned fairness* value (0-3, 4-6, 7-9), and the *absolute fairness* (the absolute value of the raw fairness value minus 5).

Both the avatar and the NPC in SVB have the same set of actions, namely *jump*, *move left* and *move right.* However, each has its own *jump height* (high or low) and *lateral speed* (slow, medium or fast) parameter value.

The two candidates for both avatar and NPC performance were *average distance to ball* (the distance a volley requires the opponent to move in order to block and return it) and *adjusted average distance to ball* (average distance divided by the opponent's lateral speed, i.e., how much time the opponent has to block and return the ball).

The only environment variable in SVB is gravity, which has two possible values (low and high).

Finally, the NPC intelligence in SVB also varies over three AI defense levels. At the lowest level, the NPC moves simply based on whether the ball is behind it or in front of it. At the medium level, it also takes into account the direction the ball is moving and whether it is falling or rising. At the high level, it calculates the spot where the ball will land.

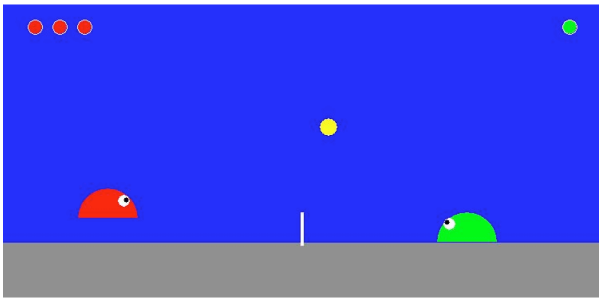


Figure 2. Slime Volleyball Screenshot

MODEL VALIDATION STUDY

As a first step toward validating the causal model in Figure 1, we conducted a study of 136 players of SVB. The basic methodology of the study was to manipulate (vary) the values at the leaf nodes of the model and to measure the values at the other nodes. Validation of the model was based primarily on direct and partial correlations. Intuitively, if factor A directly influences factor B in the model, then there should be a strong correlation between the variations in A and in B. Conversely, if there is no direct or indirect influence (causal path) in the model from A to B, then there should be little or no *partial correlation* between A and B, adjusting for the influence of the intervening variables. For example, our causal model posits that the partial correlation between difference in performance and fun, adjusting for the influence of game metrics, will be close to zero.

The top two layers of the causal model in Figure 1 have nine leaf nodes: player, ambience, perceived fairness, environment, NPC intelligence, avatar/NPC actions and avatar/NPC parameters. In this study, the ambience and avatar/NPC actions were kept constant, and we decided to measure rather than try to control perceived fairness, because we were unsure of its direct influences. This left us with five varying leaf nodes: player, environment, NPC intelligence and avatar/NPC parameters.

In addition to validating the edges in the top two layers of the causal model, a second goal of the study was to evaluate alternative instantiation candidates, both for leaf and other nodes in the third SVB-specific layer of the model. For example, we wanted to determine whether both or either of jump height and lateral speed were valid direct influences on NPC performance, and whether margin of victory, absolute margin of victory or duration were valid game metrics in terms of their direct influence on fun.

Experimental Setup

Slime Volleyball is an open source JavaScript game originally created by Quin Pendragon for two players. Among other things, we modified the game to be used by one player versus an AI-controlled NPC.

The game was made publicly available on a web page hosted on a local server. Potential participants were solicited by word of mouth and via various mailing lists and directed to the web page where, before participating, they had to read a brief introduction explaining the game and the data collection procedure. After agreeing to participate, players were asked their age, gender and what type of gamer they considered themselves to be. The breakdown of participants is shown in Tables 1.1 to 1.3.

After answering the three questions about themselves, participants were directed to the game. Each participant was assigned one of 54 possible configurations of the game (see below) according to a predetermined sequence based on order of login. After each match (five points), the participant was asked the following two questions:

- How much fun was this match? (0-9)
- How fair was this match? (0-9 where 0 means computer had advantage, 5 was an even match, and 9 means you had the advantage)

Table 1.1 Participants by Age

	Frequency	*Percent*
20 to 35	95	69.9
26 to 30	1	.7
31 plus	10	7.4
under 20	30	22.1

Table 1.2 Participants by Gender

	Frequency	*Percent*
female	46	33.8
male	90	66.2

Table 1.3 Participants by Gamer Type

	Frequency	*Percent*
casual	58	42.6
normal	35	25.7
hard core	27	19.9
unknown	16	11.8

Participants were allowed to play as many matches as they wanted (cf. duration).

The 54 possible game configurations were determined as follows: 3 AI levels × 2 gravity values × 3 jump height differences × 3 lateral speed differences. To reduce the total number of configurations, we did not use all possible combinations of avatar and NPC jump height and lateral speed parameter values. Specifically, we only used the avatar-NPC difference combinations high-low, medium-medium and low-high for each parameter.

Analysis and Results

As the starting point of the analysis, the correlation or partial correlation between *every* pair of nodes in the top two layers of causal model was calculated. We used a simple correlation for node pairs with either a direct influence or no influence between them in the model, and a partial correlation, holding the intermediate nodes constant, for nodes with a causal path between them. For each node with multiple candidate quantities in the third layer of the model, the appropriate correlation was calculated using each candidate.

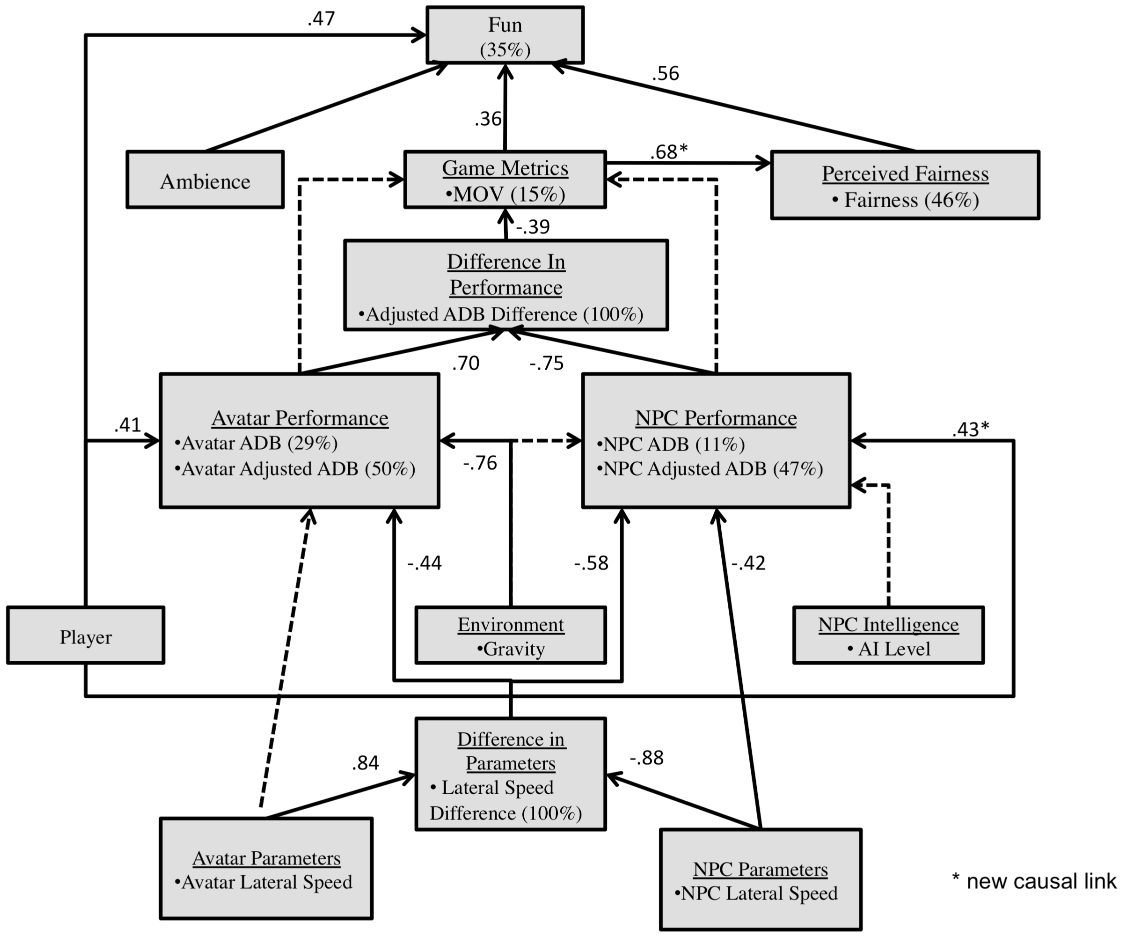


Figure 3. Validated Causal Model for Slime Volleyball

Then, to consider an edge (direct influence) in the causal model to be validated, we required that it account for at least 10% of the variance (i.e., $R^2 > 0.1$, or $R > 0.32$). Given our sample size, this restriction is more stringent than $P < 0.05$, so all of the edges in our model are statistically reliable.

Figure 3 shows the result of applying this validation procedure. The solid edges in this figure (except for the two marked with an asterisk) are the validated edges from Figure 1, with the R values indicated. For nodes with multiple instantiation candidates (avatar/NPC performance) the average correlation is reported. The five dotted edges in Figure 3 are influences in Figure 1 that were not supported by this study. The two edges marked with an asterisk are direct influences discovered in this study that were not in our original model. The direct influence of the margin of victory game metric on perceived fairness is a preliminary answer to our original uncertainty about the source of perceived fairness. The direct influence of the player on NPC performance makes sense in SVB because a strong player will cause the NPC to make more mistakes.

The percentages reported within some nodes represents the fraction of the overall variance accounted for by the current model (using ANOVA). To avoid over-fitting, we report the adjusted R^2 (a statistical adjustment to penalize for model complexity) value. For example, it appears that in SVB, our model accounts for 35% of the variance in fun.

Finally, it is interesting to note which candidate factors in the SVB-specific model turned out to be significant. For game metrics, only margin of victory was significant; for perceived fairness, only the raw fairness number; for avatar/NPC performance, only the adjusted average distance to ball; and for avatar/NPC parameters, only the lateral speed. This is useful knowledge for dynamically adjusting the difficulty of this game (see next section).

FUTURE WORK

Refining the Model

The model in Figure 1 and our pilot study with SVB are only a first step, which demonstrates the feasibility of developing a practical mathematical model for automatic balancing of computer games. To further refine the model, it needs to be applied to other game genres and specific games. In particular, further work is needed to explore the influence of ambience and the factors which influence perceived fairness. Also, to apply the model practically (see next), manipulation studies need to be performed to quantify not only the degree of correlation between factors, but also the detailed functional mapping from values of one factor to the values of another.

Applications

Returning to our original motivation in undertaking this work, we are now in a position to start using the causal model we have developed as the optimization objective function to automatically adjust parameters in a game (Andrade et al. 2005). When this adjustment is done during game play, it has been called "dynamic difficulty adjustment" (Hunicke and Chapman 2004) in recent games, such as Max Payne and Left 4 Dead. Without an explicit causal model like ours, these games have nonetheless successfully adjusted factors such as damage, aiming and spawning locations based on how the player is performing. We expect that an explicit model will make this kind of adjustment easier and more reliable.

CONCLUSION

One of the major faults of recent work on applying machine learning and optimization techniques to computer games is that it does not take the user experience into account. The causal model presented here addresses this problem by providing a mathematical model to which to which these techniques can be applied, as well as a formal framework within which to analyze and justify game design choices in general.

REFERENCES

Adams, E.; and Rollings, A. 2007. *Game Design and Development: Fundamentals of Game Design*. Pearson Prentice Hall, Upper Saddle River, NJ.

Andrade, G.; Ramalho, G.; Santana, H.; and Corruble, V. 2005. "Automatic Computer Game Balancing: A Reinforcement Learning Approach." In *Proc. 4th Int. Joint Conf. on Autonomous Agents and Multiagent Systems* (The Netherlands, July 25 - 29). ACM, New York, NY, 1111–1112.

Burns, K. 2006. "Fun in Slots." In *Proc. IEEE Symposium on Computational Intelligence in Games*, (Reno, Nevada, May 24-26).

Burns, K. 2006a. "Bayesian Beauty: On the ART of EVE and the Act of Enjoyment." In *AAAI Workshop on Computational Aesthetics* (Boston, MA, July 16).

Burns, K. 2007. "EVE's Entropy: A Formal Gauge of Fun in Games." *Advanced Intelligent Paradigms in Computer Games*, 71. Springer, Berlin, 153–173.

Cohen, P. 1995. *Empirical Methods in Artificial Intelligence*. MIT Press, Boston, MA.

Csíkszentmihályi, M. 1998. *Finding Flow: The Psychology of Engagement With Everyday Life.* Basic Books, NY.

Hefny, A.; Hatem, A.; Shalaby, M.; and Atiya, A. 2008. "Cerberus: Applying Supervised and Reinforcement Learning Techniques to Capture the Flag Games." In *Proc. 4th Artificial Intelligence and Interactive Digital Entertainment Conference* (Stanford, CA, Oct 22-24).

Hunicke, R.; and Chapman, V. 2004. "AI for Dynamic Difficulty Adjustment in Games." In *AAAI Workshop on Challenges in Game Artificial Intelligence* (San Jose, CA, July 17), 91–96.

Hunicke, R. 2005. "The case for Dynamic Difficulty Adjustment in Games." In *Proc. ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (Valencia, Spain, June 15-17,). ACM, New York, NY, 429-433.

Piselli, P. 2006. *Relating Cognitive Models of Computer Games to User Evaluations of Entertainment.* MS thesis, Dept. of Computer Science, Worcester Polytechnic Institute.

Sweetser, P.; and Wyeth, P. 2005. "GameFlow: A Model for Evaluating Player Enjoyment in Games." *Computational Entertainment,* 3, 3.

Ulam, P.; Jones, J.; and Goel, A. 2008. "Combining Model-Based Meta-Reasoning and Reinforcement Learning For Adapting Game-Playing Agents." In *Proc. 4th Artificial Intelligence and Interactive Digital Entertainment Conference* (Stanford, California, Oct. 22-24).

Yannakakis, G. N.; and Hallam, J. 2007. "Towards Optimizing Entertainment in Computer Games." *Applied Artificial Intelligence,* 21, 10, 933-971.

Jeffrey Moffett is a graduate student in the Computer Science Department at Worcester Polytechnic Institute. He graduated with a Bachelor of Science from Worcester State College. His research interests include machine learning and data mining, in particular how they apply to computer games. The work described in this paper is part of his Master's thesis.

Charles Rich is a Professor of Computer Science and a member of the Interactive Media and Game Development faculty at Worcester Polytechnic Institute. He was previously a Distinguished Research Scientist and founding member of Mitsubishi Electric Research Labs. He earned his Ph.D. at the MIT Artificial Intelligence Lab. Rich is a Fellow of the Association for the Advancement of Artificial Intelligence, and has served as chair of the 1992 International Conference on Principles of Knowledge Representation and Reasoning, cochair of the 1998 National Conference on Artificial Intelligence, and program cochair of the 2004 and general chair of the 2010 International Conference on Intelligent User Interfaces

Joseph Beck is an Assistant Professor of Computer Science and a member of the Learning Sciences faculty at Worcester Polytechnic Institute. He was previously a postdoctoral fellow in the LISTEN project at Carnegie Mellon University and a Research Professor at WPI. Beck earned his Ph.D. at the University of Massachusetts, Amherst. He was program cochair of the Workshop on Educational Data Mining at the 2006 National Conference on Artificial Intelligence.

BLENDER: EXPERIENCE OF USING RAD TOOLS FOR 3D GAMES DESIGN

Edward F. Bird, Robin Korwin-Kochanowski and Ruth Lock
University of Wolverhampton
Wulfruna Street, Wolverhampton,
UK, WV1 1LY.
E-mail: E.F.Bird@wlv.ac.uk, r.j.k2@wlv.ac.uk, Ruth.Lock@wlv.ac.uk

KEYWORDS

Game development, Blender, Prototyping, AI, Python.

ABSTRACT

The paper discusses the use of Blender in order to build a prototype 3D game. The game consists of three rooms, various objects, an animated figure and two non-player characters. Blender was used to model the 3D environment, animate the player character, apply physics to objects and impart behaviours onto non-player characters. The rooms were built separately and then integrated into a single application. The player can control the main character using the keyboard. The player character is viewed from a third-person perspective and is able to walk around the rooms and interact with the environment.

INTRODUCTION

The project was undertaken as part of a module for an MSc Computer Science Games Programming specialism. The team had a total of twelve weeks to design and implement a 3D game using an appropriate game engine.

A design for a third-person adventure game was developed. The game was entitled *Wizard Trials* and the concept was of a trainee wizard having to solve puzzles and perform tasks in order to receive a magic wand.

The team researched a number of games engines and decided that Blender should be used to create the product.

Blender was chosen chiefly because of its reputation as a complete prototyping toolset and because it was freely available. It is stated on the Blender website by Veldhuizen that “Blender is the free open source 3D content creation suite, available for all major operating systems under the GNU General Public License.”

The Blender suite consists of a 3D modelling environment, animation tools and a game engine. Thus it would be possible to develop the product with a single piece of software and would preclude the use of prohibitively expensive software such as Autodesk 3ds Max.

In addition to this, it was noted that several books had been published for Blender and that this would provide invaluable assistance as none of the team members had any previous experience of using it.

GAME DEVELOPMENT WITH THE BLENDER ENGINE

Blender’s original concept was developed in December 1993 but the game engine was not added until eight years later in 2001 v2.10. It was written by Erwin Coumans (main architect) and Gino van den Bergen (collision detection and physics). Version 2.48a was the latest version being used for this particular project.

Blender allows for the creation of 2D and 3D, models, UV mapping, texturing, rigging for character animation, skinning, animation, particle simulation, rendering and games creation, with the inclusion of a fully functioning physics engine to support game development which includes collision detection. It also uses a scripting language called Python to allow for more complex program development for games play.

Various games and movies have previously been created using Blender.

Elephant Dreams (2005) is a 10 minute short open movie commissioned by Netherlands Media Art Institute which shows two people who explore a strange mechanical world, with each having their own perception of what is actually there.

Then there was another 10 minute open short movie entitled *Big Buck Bunny* (2007) produced by the Blender institute in Amsterdam which depicts a giant rabbit finding his happy sunny morning walking being disturbed by three rascally rodents.

In 2008 development then progressed to a fully-functional game called *Yo Frankie*. This was again supported by the Blender institute in Amsterdam but was a truly international development team with members from Australia, USA, Poland, Germany, Spain and Argentina. The game involves the player controlling an evil rodent Frankie, who explores the forest seeking other animals to harass. It is free to download at http://www.yofrankie.org/.

Blender was created by professional graphics designers for professional graphics designers. Therefore it has a very strong visual interface with lots of functionality built in allowing the user to manipulate graphics objects in a myriad of ways and also allowing for the interface to be fully customizable.

A user can split the screen into several working areas each with a different view of the model being worked upon. This helped support the different working methods of all 3 developers.

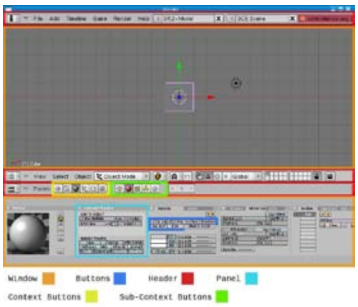


Figure 1: Blender Environment

Even though there are toolbars and menus within Blender the preferred method of working is using shortcut keys for speed and ease of use. There are at least 100 shortcut keys which once learned makes development very fast. However for a novice it makes the environment challenging to learn. All three developers kept a crib sheet containing all the relevant shortcut keys.

Modelling

Modelling of all three rooms was facilitated using only a few of the many objects that were available within Blender. The most common objects used were the mesh primitives called the cube and the UV sphere. The cube was used to build up the architecture for each individual room creating the walls and floors in the hall, drawing room and library.

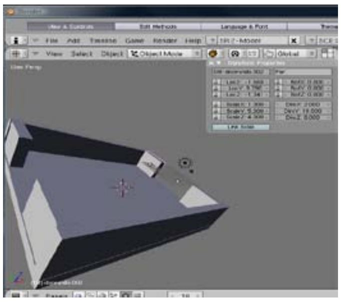


Figure 2: Creating Room Architecture

Once all of the rooms had been created it was then necessary to colour the objects so that they appeared to be more realistic. Various techniques were tried including using materials, textures and UV mapping using external image sources but due to the wizard character requiring collision detection it was deemed necessary that all objects within each room would have to be UV mapped.

UV Mapping

UV mapping is a technique used to wrap a 2D image or image texture such as a .jpg image onto a 3D mesh such as a cube. U and V are the axes of the planes which represent X, Y and Z coordinates in 3D space.

The walls and floors of each room were created using large individual cubes stretched out to create a solid surface. When trying to map an image tile onto such a large area it was found that the image became distorted and stretched to fit the surface area. Each wall and floor had to be subdivided up using a technique called looping to increase the number of faces and then the image was mapped to each face. A clear stone tile was used to map the walls of each room with a wooden .jpg tile used to map out the floors.

Another difficulty which was not realised until later was that when a .jpg image file is mapped to an object, an option called relative path had to be set when selecting the image. If this was not chosen then any changes to the file path would cause the UV mapping to be lost because the Blender file would no longer be connected to the correct location of the image. This was essential when integrating the rooms to create the final product.

Game Physics

The Blender Game Engine (BGE) has a built-in physics engine called Bullet. This handles collisions between objects in a scene and allows for forces such as gravity to be applied. In the drawing room scene, a sphere representing a boulder should roll around the room and hopefully fall into the well. In order to achieve this effect the sphere was converted into a dynamic object. By setting it as a dynamic object the Blender game engine recognises it as a physical object. 'Rigid body' was also selected enabling the object to rotate correctly.

Figure 3: The Drawing Room

Logic Bricks

The BGE provides a visual tool for programming object behaviour. This is possibly the strongest recommendation for using Blender as a game prototyping tool. Strong similarities exist between the game engine design and behaviour-based architecture (Nakamura 2006), a model for designing the behaviour of robotics agents. Information gathered from the environment through sensors is processed and results in actions performed by actuators.

In BGE, object behaviour is defined in terms of groups of sensors connected to logic (or scripted) controllers, which in turn are connected to actuators. Blender's logic buttons enable sensors, controllers and actuators to be simply added and connected.

Sensors include collision, keyboard, mouse and joystick to name a few. There is also an *always* sensor, which fires continually (or just once). Sensor output is a Boolean value.

Controllers can perform a simple logic test on their sensor connections (AND, OR XOR, etc) and activate their actuators if a true value is returned. They can also contain simple expressions or a link to a Python script.

Actuators can create motion, change states (see below), play animations and many other actions.

In the Wizard Trials prototype we used logic bricks in all three rooms although the hall (Towers of Hanoi) was still in early development at the end of the project. The main areas requiring substantial logic development were:

- The drawing room floor
- The wizard
- The library (sentinels)

Keyboard sensors were used to enable the player to tilt the drawing room floor by pressing the cursor keys. The purpose of this was to roll a large boulder in the room into a hole.

The floor was given a series of logic bricks each with a keyboard sensor, a simple *AND* controller and a motion sensor to rotate it along the local X or Y axis. Figure 4 shows a sensor-controller-actuator chain for the up key, which rotates the floor in a positive direction around its Y axis.

Figure 4: Logic Bricks for the Up Key

The hole in the floor contains a tiny cube which is used as a collision sensor – when the boulder touches it, the room can be reset.

Blender allows you to add properties to objects; these can be accessed using logic bricks or Python scripts. A (real-time) timer property stores the elapsed time since the game began. This is used in the drawing room to limit the time the player is allowed to complete the task.

We adapted the wizard from a figure included in a Blender character animation tutorial (McKay 2008). This player character was useful in that it had a walk cycle that needed only minor alteration for our purposes.

The wizard's logic bricks can be separated into three groups: those which enable animation, those which add movement and a Python script for publicising the object's location and calculating various hit points, magic points (etc).

An *always* sensor is used to trigger the character's idle cycle and Python script. This gives the character a default action.

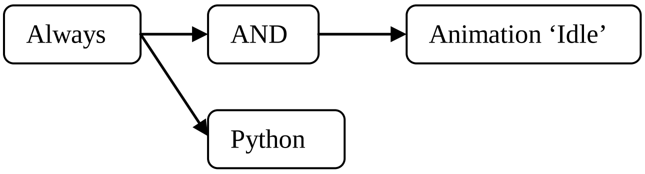


Figure 5: The Idle Cycle and Python Script

It is necessary to 'publicise' the wizard's position because problems had been encountered when trying to access the real-time position of the wizard from other objects (i.e. the sentinels in the library). Blender is object oriented and game objects do include a *getPosition()* method. However it was found that accessing this from another object, using say, *Wizard.getPosition()*, returned the Wizard's static initial position. Whereas the Wizard did have access to its current position with *me.getPosition()*.

The work-around (if indeed, it was one) was to have the wizard copy its position to global variables every game cycle. Of course this is not efficient and at the very least, could have been greatly improved by only running the script when the wizard moves.

An interesting challenge was encountered while adding motion to the wizard. At first, this was facilitated by setting the dynamic Y location in the motion actuator to a negative number (effecting forward movement by a series of tiny jumps – one per frame). This caused a problem in that collision detection did not work properly and the character was able to pass through walls.

A partial solution was found by utilising the physics engine: applying a force to the figure (effectively pushing it along). This worked in that if the force was small enough, the character would collide with walls (rather than passing through them). However, the motion was uneven and the character would seem to slide to a halt or accelerate greatly if the forward key was held down. The problem was solved by replacing force with linear velocity in the motion actuator.

Blender's motion actuator offers a range of methods for creating and fine-tuning movement. This complexity can be confusing at first but is characteristic of the games engine's flexibility and power.

The State System

The provision of states adds another dimension to the logic bricks. A state is a group of related logic bricks and is often linked to a certain activity or 'state' requiring a particular behaviour. An object can have a number of active states at any given time.

States were used in several places in the Wizard Trials game. A simple example would be the drawing room where once the room is entered, the floor activates in state 1, which contains logic to allow it to be tilted; when time is up, the floor resets and changes to state 2 where it cannot be moved.


Figure 6: The Library

In the library, there are two sentinels roaming the maze trying to catch the wizard.

The sentinels make use of the state system to implement their behaviours. The Sentinel states are as follows:

1. Initialisation – The sentinel is moved to a random place (but complementary to that of its counterpart).
2. Pursue – Head towards the wizard.
3. Attack – Attack the wizard.
4. Hide – Go to a good hiding place.
5. Random – Go to a random place then repeat.

There are two more states, which are always active:

6. Game AI – Check current circumstances and switch states if necessary
7. Set Active – Enable or disable the sentinels as necessary (Check the player's status and position).

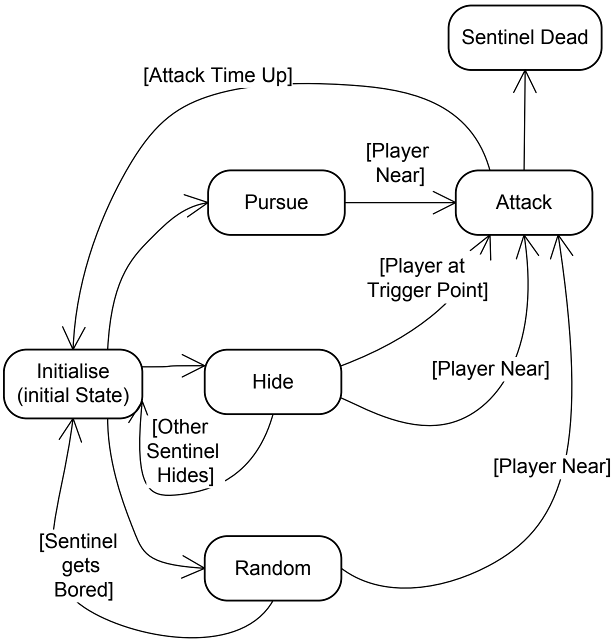


Figure 7: The Sentinel Finite State Machine

Each of the first five states consists of an initialise (once) and a general (always) script, which utilise a waypoints algorithm to move the sentinel to an appropriate place.

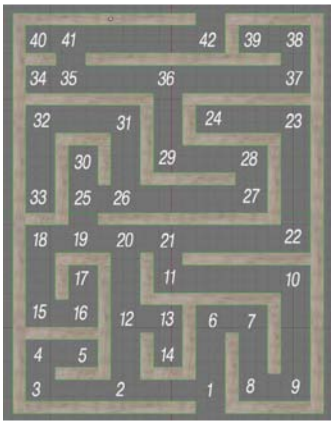


Figure 8: Library Waypoints

The sentinels track the wizard and their own position as they pass over waypoint nodes. It is then possible using a table to gain the next node in the path from the sentinel to the wizard (or a hiding place etc).

```
def setAIState():
   wizX = GameLogic.globalDict["pcX"]
   wizY = GameLogic.globalDict["pcY"]

   if me.active:
        if me.ai_state == 1:          # Initialise
             restartSentinel(True)

        elif me.ai_state == 2:        # Pursue
             if cont.getSensor("nearTarget").isPositive():
                  me.check_point = me.game_time
                  # Attack
                  jumpToState(3)
             else:
                  return

        elif me.ai_state == 3:        # Attack
             if me.game_time - me.check_point >= 6.0:
                  # Tidy up position - if not at a node, continue to the next
                  goOnToNode(me.getPosition()[0],  me.getPosition()[1])
                  # Teleport to random start point and restart
                  restartSentinel(False)

        elif me.ai_state == 4:        # Hide
             # if in hiding place and PC on a trigger node
             # or not yet hidden (but at a node) and PC at next node
             if (me.hidden and (me.target_node == me.trigger_node or \
                       me.target_node == me.trigger_node2)) or \
                       (not me.hidden and me.at_node and wizAdjacent()):
                  me.hidden = False
                  # Pursue
                  jumpToState(2)
             elif otherState() == 4:
                  restartSentinel(True)
             else:
                  return

        elif me.ai_state == 5:        # Random
             if wizAdjacent():
                  jumpToState(2)
             elif me.bored == 0:
                  restartSentinel(True)
             else:
                  return
```

Figure 9: Python Function Controlling Sentinel States

Figure 9 shows the state control function from the AI Python script. This function is called every game cycle (for each sentinel). In AI State 2 (Pursue) the script checks the state of a 'Near' sensor, which is true if the wizard is 1.5 units (representing metres) away from the sentinel. If it is triggered then the sentinel switches to attack mode (State 3). Blender Python scripts are able to easily manipulate logic bricks [jumpToState(n) merely sets up and triggers a State actuator].

It is worth noting that the sentinel state system requires quite a lot of scripting to implement and possibly this could begin to degrade performance in a larger more complex game. However, the prototype we produced did work well.

The BGE logic bricks and state system are powerful and simple to use but do require scripting to implement more complex game AI. There is no inbuilt AI engine with Blender.

EVALUATION AND CONCLUSIONS

Given the time constraints (seven weeks at most for implementation) and the group's lack of experience with Blender, the prototype was expected to be in some ways incomplete. In this section we will consider issues that arose in development and some aspects of the BGE, which are pertinent to RAD development of games.

Discussion

Several features of the game required the use of particle systems. Examples include the bubbling, steaming cauldron in the hall, a fire in the drawing room and visible magic attacks for the wizard and sentinels. We were unable to discover how to include real-time particle effects in a game and simply ran out of time. Much of Blender's supporting literature (official and otherwise) focuses on particle effects for animation rather than games.

Blender is well equipped to enable group projects. A game may comprise many blend files, which can be appended (one file copied into the other) or linked (by reference only). Linking allows separate development of objects or properties (such as mesh, material etc). It is possible to group multiple objects and properties for linking.

The prototype was subdivided into rooms, objects, the sentinels and the wizard. This enabled us to each work on separate rooms independently and kept the blend file sizes to a minimum. The issue of file size was important because Blender appeared to become unstable with overlarge files.

Blender facilitates the creation of multiple scenes. However, due to inexperience, we failed to make use of this feature and created the prototype as a single scene. This caused several problems in developing the rooms: which needed to be rotated and aligned at integration. It was also necessary to write scripts to disable rooms when the wizard was not present.

A better strategy would have been to place each room in a separate scene. Blender provides for the creation of scene portals and a scene can be easily deactivated and reactivated (without the need for scripting). Blender does also enable the use of blend file portals however the method of transfer of camera, player and data between files was too complicated to achieve within this project.

Blender's debugging facilities are somewhat limited. A second, console window is used to output information such as warnings, error messages (etc); it can also be used to *print* values from Python scripts. This enables variable and flow checking within a script but does not easily afford such testing/debugging for logic bricks. At times, it was necessary to resort to replacing logic bricks with scripts unnecessarily because this was the only way to debug the operations.

Debugging was necessarily quite primitive. For example: early in the sentinels' development, the AI State function (Figure 9) was not working properly: the sentinels did not seem to be communicating their states to each other (see state 4). Print statements were inserted to display the sentinel's number, state and perceived "other sentinel's state". This did help resolve the problem but break points and stepping facilities would have greatly simplified the process.

Figure 10: The Camera Actuator

Blender's Camera actuator enables quick and easy development of an orbiting camera. Figure 10 shows the actuator set up to follow the wizard (grp.wiz) , keeping 3 units above and at a distance of 3-4 units; the camera will try to remain behind the wizard's Y-axis. This provides a good starting point but further work is required to prevent the camera from passing through walls (etc). Blender does include ray casting, which can be used to prevent the target from becoming obscured (Barton, 2008). A ray cast from the target is sensed by the camera; should any object obscure the ray, the camera can move to restore line-of-sight.

Given the time constraints, it was not possible to implement ray casting for the camera in this prototype. Two approaches were used: while in the open space of the Hall and Drawing Room, the camera uses simple orbiting; as the wizard enters the confines of the Library maze, the camera switches state to enable a Python script controlled hover mode (position fixed above the wizard looking down).

Figure 11 shows the script that switches the camera state as the wizard enters/exits the Library. The line containing *cont.getActuator("Activate")* gets the state actuator from the camera's logic bricks (*cont* is assigned to 'this controller' higher in the script). The state change operation is set to 0 for a copy operation (to copy the new mask replacing the old). Blender's states are activated using a 30 bit binary number (30 states) so the mask 34 (binary 100010) turns on states 2 and 6. Different operations allow binary logic to be used in applying the mask; this enables states to be easily added, removed or toggled. Finally, the active state actuator is *added* and the operation performed

The parameter Boolean 'free' is given a simple expression in the function call, which is true if the wizard is not in the Library.

```
#Set cam state
def setCam(free):
  myState = me.getState()
  if free and myState  == 33:
      act = cont.getActuator("Activate")
      act.setOperation(0)                 # copy
      act.setMask(34)                     # orbit
      GameLogic.addActiveActuator(act, 1)
  elif not free and myState  != 33:
      act = cont.getActuator("Activate")
      act.setOperation(0)                 # copy
      act.setMask(33)                     # look down
      GameLogic.addActiveActuator(act, 1)
```

Figure 11: Python Script Controlling Camera States

Figure 12 shows the script for moving and rotating the camera in the Library. Firstly the camera is positioned above the wizard at a height of 4.5 units. The wizard's rotation matrix is copied from the global dictionary (where the wizard updates it every frame). It is necessary to rotate the camera 180° around the Z-axis (otherwise the camera appears to face backwards). This is done by negating the X-x, X-y, Y-x and Y-y values.

```
wizX = GameLogic.globalDict["pcX"]
wizY = GameLogic.globalDict["pcY"]
me.setPosition([wizX, wizY, 4.5])
camRot = GameLogic.globalDict["pcRot"]
camRot[0][0] = 0 - camRot[0][0]
camRot[0][1] = 0 - camRot[0][1]
camRot[1][0] = 0 - camRot[1][0]
camRot[1][1] = 0 - camRot[1][1]
me.setOrientation(camRot)
```

Figure 12: Python Script to Position the Camera

Conclusions

The Blender integrated development environment and games engine does provide the tools necessary for the rapid creation of prototype games. Its emphasis on the use of mouse and hotkeys does in our experience make it hard to learn but increases productivity once it is learnt.

We found that the 3D development environment enabled very rapid creation and texturing of objects (although it is noted that our requirements were minimal).

The Logic Brick system does facilitate speedy addition of game logic to objects. Python scripting enables more complex features such as Game AI. The combination of Logic Bricks, States and Python scripts can make debugging complex; this being exacerbated by the limited toolset.

It is accepted that while we were unable to fully implement features such as advanced camera operation, shadows, real-time particle systems and blend file portals, these problems were a reflection of our inexperience and time constraints rather than Blender's features.

Blender provides an impressive array of tools, which once learnt, enable rapid development of prototypes and games. Provision for group collaboration in the development process is built into the program.

REFERENCES

Barton, C. 2008. "Yo Frankie! The Open Game Project." In The Blender GameKit, C. Wartmann and M. Kauppi eds. Netherlands: Blender Foundation. 159-169.

Brito, A. (2008) "Blender 3D: Architecture, Buildings, and Scenery", Packt Publishing.

Gumster, J.(2009) "Blender for Dummies", Wiley Publishing.

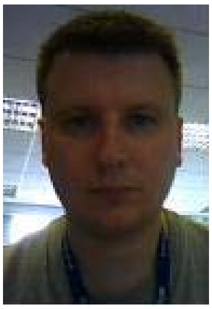
McKay, R. 2008. "Game Character Animation Using Armatures." In The Blender GameKit, C. Wartmann, and M. Kauppi, eds. Netherlands: Blender Foundation, 129-147.

Mullen, T.(2008) "Bounce, Tumble, and Splash" Wiley Publishing, Inc.

Nakamura, T. 2006. "The Soprants: conceptual and technical framework for a 3D interactive video game." M.Sc. Thesis, Texas A&M University, 19-20.

Veldhuizen, B. 2009. "About Blender". Available: http://www.blender.org/. Last accessed 28 May 2009.

BIOGRAPHY



Edward Bird graduated with honours in Computer Studies in 2002 and has since worked as a software developer and lecturer. He is currently completing an MSc in Computer Science (Games Development).

Robin Korwin-Kochanowski graduated with honours in Computer Science in 1991. He majored in Artificial Intelligence and Real-Time Systems. Serious health problems put his career on hold for a number of years. He has been supporting people with dyslexia in the use of voice recognition software for ten years at City of Wolverhampton College.

After improvements in his health, he has resumed his studies and is now in the final stages of an MSc in Computer Science (Games Development).

Ruth Lock graduated with honours in Computer Science in 1996. She completed a PGCE in FAHE in 1999 and for the last 15 years has been teaching programming, database design, project management, HCI and information systems to HND and Foundation Degree students at the City of Wolverhampton College.

She recently resumed her studies and is now in the final stages of completing an MSc in Games Development studying at the University of Wolverhampton.

THE USE OF NARRATIVE IN THE DESIGN OF SERIOUS GAMES FOR CRIMINAL INVESTIGATION AND COMMUNITY POLICING

Daniel Andrews and Chris Baber
The University of Birmingham
Dept. of Electronic, Electrical and Computer Engineering
Birmingham, UK
E-mail: dxa847@bham.ac.uk, c.baber@bham.ac.uk

KEYWORDS
Serious Games, Interactive Narrative, Crime Scene Investigation, Community Policing

ABSTRACT

The use of storytelling and narrative in commercial videogames has its recognized advantages, but not in Serious Games to the same extent. By demonstrating the use of narrative in both crime scene investigation and community policing, it is hoped that the usefulness of narrative in the realm of Serious Games can be demonstrated. This is because narrative is a convenient way to design realistic characters and perspectives within a range of scenarios, as well as the scenarios themselves. If adequately researched the resulting game would be multidisciplinary and realistic, providing a comprehensive training experience.

INTRODUCTION

The use of storytelling and narrative in commercial videogames can be of benefit to many aspects of entertaining gameplay (Adams and Rollings 2007, ch. 7). Yet the potential benefits of this in Serious Games could be less apparent. Serious Games can be defined as games that educate, train, and inform (Michael and Chen 2006). Developers of Serious Games might therefore perceive storytelling or narrative to be redundant qualities. However, in his analysis of the learning principles of videogames, James Paul Gee (2007) has argued that many of the design features of videogames contain a great many principles associated with good learning, including storytelling and narrative. A Serious Game that makes use of these aspects could therefore provide a compelling and effective learning experience.

Our research is exploring the potential use of narrative and storytelling to inform the design and development of effective Serious Games. Our approach so far has been to identify disciplines in which narrative is a fundamental, if not immediately obvious, aspect of their functioning. This research is primarily focused upon the realm of criminal investigation. However, as we will attempt to demonstrate in this paper, storytelling and narrative in Serious Games can be a useful tool in the convergence of seemingly different but related disciplines in order to provide comprehensive learning scenarios.

To demonstrate this we will explore two distinct aspects of criminal investigation:

1) Crime Scene Investigation, and
2) Community Policing

In exploring these two areas separately we will propose a design for a Serious Game that will effectively train officers in both disciplines, without compromising the information learned in either.

CRIME SCENE INVESTIGATION

One of the most immediate associations with Crime Scene Investigation in the realm of narrative and storytelling would be the television series of the same name. It could be argued that the popular fixation with the sub-genre of forensic fiction dates back at least to the emergence *Sherlock Holmes*, which demonstrates the extent to which this discipline has been adopted for its entertaining narratives. It could therefore be the inherent narrative qualities of criminal investigation that result in it being adopted by media producers. Baber et al. (2006) have shown that a narrative model tends to be adopted in crime scene investigation in order to make sense of evidence. In volume crime this is particularly important, as there are limited resources to examine the scene, and the most probable narrative of what occurred must be identified in order to quickly secure the most relevant evidence. There are, of course, inherent problems in such an approach, as the identification of evidence can be influenced by the presumptions of the investigator. In order to avoid this the UK's Forensic Science Service recommends that crime scene investigators adopt at least two mutually exclusive perspectives when examining evidence, which would normally correspond to the perspectives of the prosecution and the defence in an oppositional model of criminal justice (Caddy and Cobb, 2004). These perspectives would essentially equate to being different interpretations, or narratives, about what happened at the scene. In addition to the evidence, different people involved in any individual crime could potentially provide an assortment of narratives, and it is the responsibility of the investigator(s) to assess all the information available to determine what 'actually' happened. These narrative shifts that can occur during complex investigations are what create dramatic tension and interest, which is why criminal investigation provides a constant source of inspiration to popular fiction.

Both the television series *CSI* and the literary investigator *Sherlock Holmes* have been adapted into videogames, which demonstrates how the potential for this genre has also been recognized in interactive media. However, in the commercial realm such adaptations tend to adopt a linear approach to narrative, which could be because of the cost and difficulty associated with developing more complex, branching narratives in videogames (Adams and Rollings 2007, ch. 7). Indeed, this was the approach used in developing a prototype for a Serious Game in *Flash* designed to train users in processing a crime scene and in bloodstain pattern analysis (Andrews, 2008). The game was an adaptation of a section of Jeffrey Deaver's novel *The Devil's Teardrop*, which meant that the characters and scenario had already been partly defined. The characters from the novel were developed into agents in the game, allowing the player to participate as the protagonist and interact with other characters in order to obtain information. The limitations of the program meant that the evidence and bloodstains had to be processed in a specific order, and unless the player answered the questions related to the point of origin of the bloodstains correctly then they could not progress through the story (see Figure 1). In this case, the linearity of the story imposed limits upon the interactivity of the game. The player had limited interaction with both the environment and the evidence, meaning the gameplay and its training capabilities were limited as a result. However, as a prototype it was able to demonstrate how narrative can be used to combine multiple facets of training within a single Serious Game. If the individual elements of the prototype were more developed in terms of their accuracy and interactivity, allowing the player to more freely and thoroughly examine a crime scene, interact with other characters, and analyse the available evidence, then a more realistic training environment would be constructed.

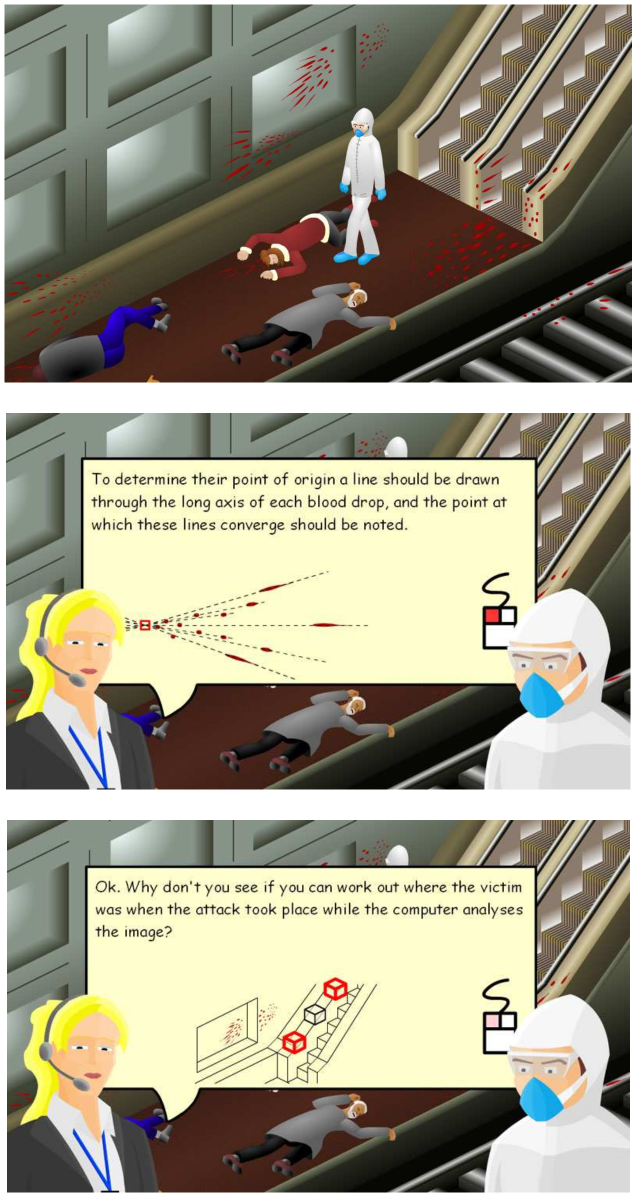


Figure 1: Screenshots from the Serious Game Prototype of Jeffrey Deaver's 'The Devil's Teardrop' – Navigating a crime scene, analysing bloodstains and answering multiple choice questions (Andrews, 2008)

In the same way that narrative has been used in the prototyping of this game, it could also be used in the more accurate representation of a crime scene. As indicated above, the evaluation of multiple perspectives, or narratives, is already an important aspect of criminal investigation, as it allows for more information to be gained and a more accurate and objective account to be realized. In order to create such an environment, several narratives must be designed in order that they correspond with the scenario. To help inform this approach we will now explore the next discipline.

COMMUNITY POLICING IN THE CONTEXT OF COUNTER-TERRORISM

In 2008 a research team from the University of Birmingham, UK, conducted a study into Police-Muslim Engagement and Partnerships for the Purposes of Counter-Terrorism (Spalek et al. 2008). The research showed that the different forms of engagement adopted by police can have repercussions that might not initially be realized by the officers themselves. Through numerous interviews with members of Muslim communities, the Muslim Safety Forum and police officers from the Muslim Contact Unit in London, the team were able to assess numerous individual narratives as representations of actual experiences.

The research conducted showed that there are many factors involved in how communities and groups respond to police officers. In conclusion it was argued that a ‘hard’ approach to practicing intelligence gathering by police can result in disengagement by communities, which can result in long-term losses in intelligence, whilst a ‘soft’ approach can result in long-term relationships, and possibly partnerships, that could be beneficial to countering violent extremism.

In terms of training, one of the most obvious ways in which these findings can be communicated to officers is via presentations, seminars, or workshops related to the report. However, upon examination of the study, a number of variables in the report have been identified which would affect engagement between officers and communities, based upon the individual narratives analysed. A Serious Game could use programming techniques in order to illustrate the complex relationships between these variables and allow for a more interactive training environment to be constructed. Figure 2 shows stills from a simulation currently being developed in *NetLogo* that adapts the qualitative results obtained by Spalek et al. (2008) into qualitative data for the purposes of developing a Serious Game. The results of the simulation depict what the report concludes, which is that a ‘hard’ approach could ultimately lead to drops in intelligence gathering over time (see graph at the bottom of Figure 2).

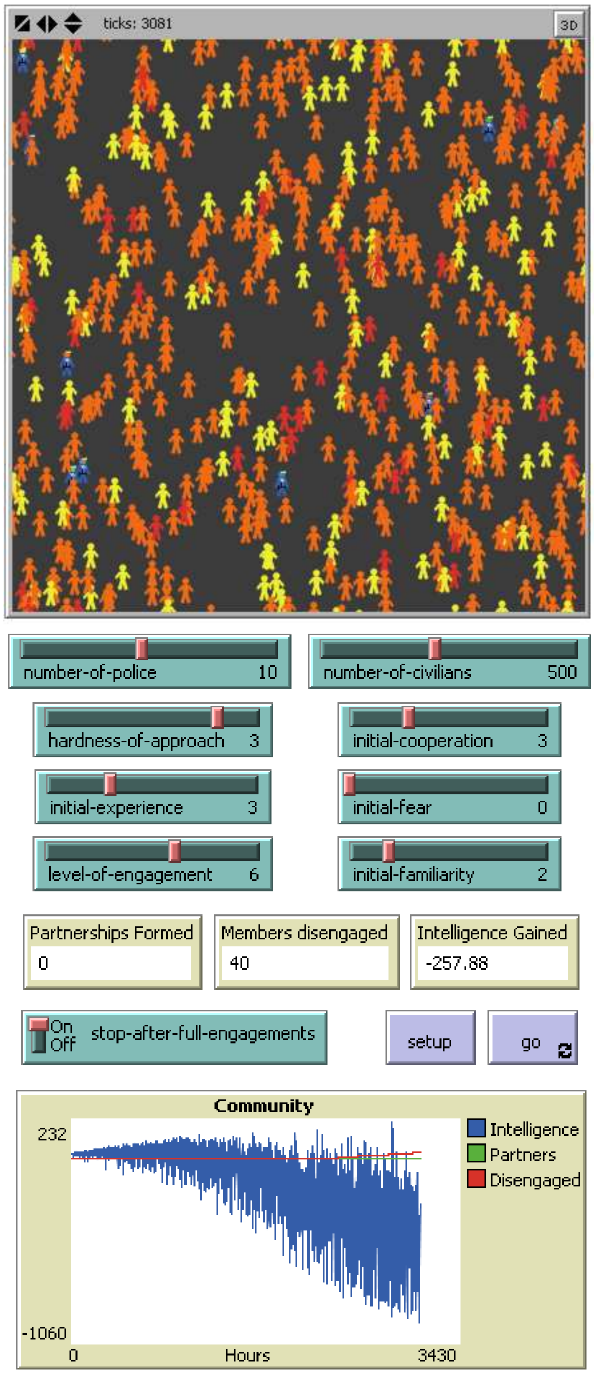


Figure 2: *NetLogo* simulation of police-community relations based upon results obtained from Spalek et al’s (2008) report

One recognised weakness in this approach is that the formulae and algorithms used in the simulation can only ever be an approximation of what actually happens, as it is impossible to accurately identify and define all the available variables, or quantify the relationships between them. Another weakness is that, in its current form, the simulation does not allow for variables to be altered within individuals or groups smaller than 'police' or 'civilians'. However, through regular referrals between all the related parties during various stages of development, it should be possible to accurately reproduce trends identified in the report to the satisfaction of everyone. The variables could, in turn, be used to inform individual narratives that reflect the responses obtained in the study by Spalek et al. (2008). Once this has been achieved, the variables can be used in a Serious Game to accurately devise the relationships between characters within a community.

DESIGNING A SERIOUS GAME IN CRIME SCENE INVESTIGATION AND COMMUNITY POLICING

As the individual values of all its variables can be independently altered, the simulation could provide an efficient way to demonstrate various relationships and present the range of findings in Spalek et al's (2008) report. In its current form the ability to interact is limited to the control of these variables and the observation of the resulting interactions. On the other hand, if the player could participate from the perspective of a single agent from the simulation the range of potential scenarios would be vast due to the amount of possible combinations between the variables. This becomes even more complicated if the variables can be independently altered between smaller groups or individuals, and the player's options would have to be kept to an absolute minimum in order to make it realistically programmable. However, if the simulation was used to help define a small amount of individuals within a specific area then a greater range of interactions could be permitted to the player without compromising the training provided or realism attained. It is therefore proposed that a Serious Game would benefit from an approach that combines the interactive capabilities of both the simulation and the prototype game described above.

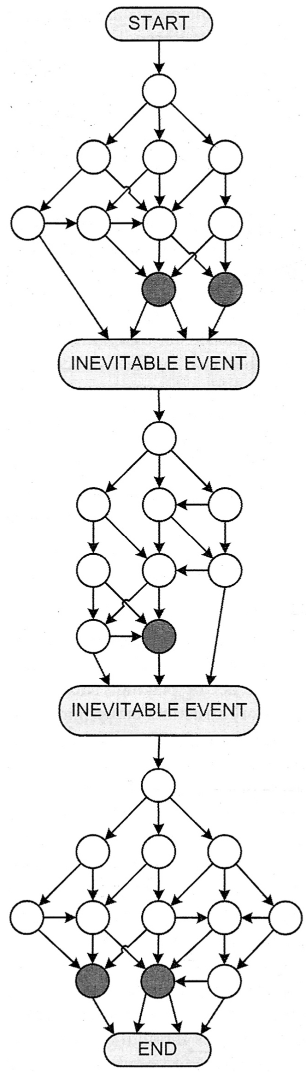


Figure 3: Simplified structure of a foldback story – the circles represent 'branch points' whilst dark spots represent branch points that always lead to an 'inevitable event' (Adams and Rollings 2007, p. 201)

Figure 3 shows an approach to narrative described by Adams and Rollings (2007, ch. 7) as 'foldback Stories', which provide a compromise between branching and linear stories comparable to the development of the proposed scenario. Although this model is presented in relation to commercial videogames, it could be used to design effective Serious Games as well. In the illustration shown, the player can take various paths between events which the game's 'Storytelling Engine' has determined must occur in order for the game to progress (Adams and Rollings 2007, ch. 7). In the development of Serious Games, such events could only occur if a combination of certain tasks have been completed or information is obtained. Whilst this makes the process of designing such a game more complicated, it nevertheless provides a more realistic approach than a linear one, without compromising the players' interactive capabilities. The proposed scenario combines elements of community policing with crime scene investigation by allowing the player to investigate a volume crime within a community setting; such as a theft from a community centre. The player would have the option to investigate the crime via a number of 'paths' in a manner of their choosing. However, a combination of approaches would be required in order to progress through the game. In accordance with how a crime would normally be investigated, the player would be required to interact with community members as well as process the available evidence in order to obtain enough information to finish the game. It would be up to the player to successfully engage with the community in order to make sense of the evidence. The players' interactions would consequently affect the outcome of the game, and inappropriate engagement decisions made by the player would make the crime more difficult to solve.

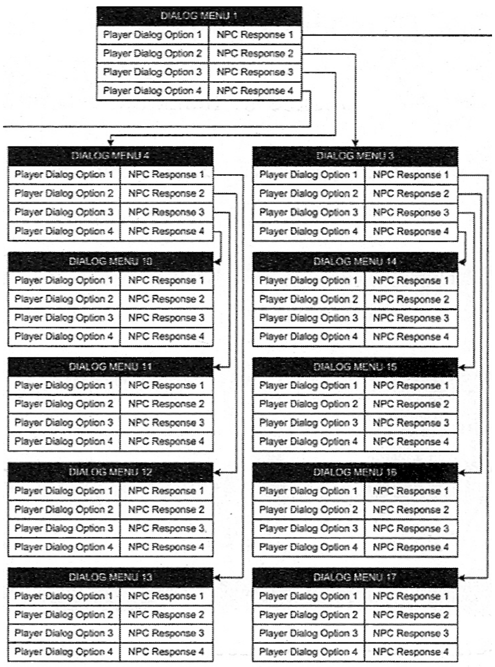


Figure 4: A simplified diagram of part of a dialogue tree (Adams and Rollings 2007, p. 212)

Figure 4 shows a small section of a dialogue tree, which Adams and Rollings (2007, ch. 7) have shown can be used to contribute to both the story and the actual gaming mechanics of a videogame. In the proposed scenario for a Serious Game, therefore, the branch points in the foldback story structure (Figure 3) would represent sections of the game involving either the analysis of evidence or the interaction with community members, whilst dialogue trees (Figure 4) could be used to define the dialogue that takes place. If the processing of evidence and the interaction with community members is adequately researched then the player should be able to reflect upon and learn from both experiences.

CONCLUSION

Established principles of designing narrative in commercial videogames could be effectively applied to a Serious Game, as long as the content is adequately researched and the player is given sufficient interactive capabilities. These design principles lend themselves well to the design of multidisciplinary training environments due to the diverse range of narratives they accommodate. Simulations based upon established research could therefore inform the design of a Serious Game, particularly if it relates to social behaviour. It is argued that this approach could provide a comprehensive training environment as well as an entertaining learning experience. Whether this argument is valid will become apparent once the prototype has been developed and evaluated.

REFERENCES

Adams, E. and A. Rollings. 2007. *Game Design and Development: Fundamentals of Game Design*. Pearson Prentice Hall, N.J.

Andrews, D. 2008 *Jeffrey Deaver's 'The Devil's Teardrop': Adapting a Novel into a Serious Game*, MSc Thesis, Nottingham Trent University, UK

Baber, C., P. Smith, J. Cross, J. Hunter, R. McMaster. 2006. "Crime scene investigation as distributed cognition." *Pragmatics & Cognition*, 14:2, 357-385

Caddy, B. and P. Cobb. 2004. "Forensic Science" In *Crime Scene to Court: The Essentials of Forensic Science*, P.C. White, ed. The Royal Society of Chemistry, Cambridge, 1-20

Gee, J. P. 2007. *What Video Games Have To Teach Us About Learning And Literacy*. Palgrave Macmillan, N.Y.

Michael, D. and S. Chen. 2006. *Serious Games: Games That Educate, Train, and Inform*. Thomson Course Technology PTR, Boston, MA.

Spalek, B., S.E. Awa, and L.Z. McDonald (2008) *Police-Muslim Engagement for the Purposes of Counter-Terrorism: an examination*, http://www.ahrc.ac.uk/News/Latest/Documents/RaR%20Islam%20Summary%20Report.pdf, accessed at 14th June 2009

PLAYER MANIPULATION

Zi Xu, Siew and Alexander Nareyek
Department of Electrical and Computer Engineering
National University of Singapore
4 Engineering Drive 3
Singapore
E-mail: u0508690@nus.edu.sg, elean@nus.edu.sg

KEYWORDS
Player Actions and Choices, Manipulation Techniques, Psychology

ABSTRACT

Video games are capable of giving the players a better entertainment experience due to the interactivity as compared to passive media like movies and books. The overall quality of the entertainment experience is highly dependent on the quality of the game content. Given that video game developers are only able to create a finite set of good quality content for the player to experience and that content that is not experienced does not contribute to the player's entertainment experience, there is a problem where game developers have to direct the player towards game content and meaningful gaming experiences while not sacrificing the interactivity of the game. In this paper, we propose solutions to this problem by studying manipulation concepts in other areas, classify the concepts into a portfolio of manipulation techniques that can be used in games, and conduct some preliminary experiments, which indicate that an implementation of the techniques is not as simple as one would imagine.

INTRODUCTION

The term "manipulation" used in this context means to influence persons to do what we want them to do. The type of manipulation that we are interested here is *subconscious* manipulation, i.e., the target should not consciously realize that any influencing is taking place. In other words, we don't want the player to consciously know that they are being manipulated. When we use manipulation techniques, we want to exploit psychological techniques in a way such that, for example, a compelling reason is created for a player that makes the player select a particular response towards a particular situation.

As such, the manipulation techniques found in this paper are derived from psychological techniques practiced in areas like marketing and retail. These techniques rely on the fact that we usually develop a preprogrammed set of actions for ourselves for a particular situation so that we can apply these actions without having to think through the situation (Cialdini 2001). For example, we usually extend our right hand to shake hands with people we meet for the first time. This action is driven by the need to appear courteous in front of the person and is automatically triggered as the most appropriate action to take when faced with this situation. Using the psychological techniques found in other literature as a basis, manipulation techniques that can be used in the gaming context are developed.

In the next section, common in-game actions are classified into categories that will be used to discuss the application scope of the manipulation techniques. This is followed by the details of the developed manipulation techniques.

CLASSIFICATION OF PLAYER ACTIONS

Guiding the player to make decisions in a certain way requires distinguishing different actions from one another. Some techniques are more effective in influencing certain types of actions than for others. The actions that a player might make vary from game to game depending on the genre as well as what type of game it is, and it is impossible to list all of the possible actions exhaustively. Instead, actions that are found commonly in games will be grouped together under a category based on the game design needs. This classification will serve as the basis of categorization of the example implementations for each manipulation technique later on.

Movement

This category includes movement actions that the player can make, for example, moving the player's avatar around in the game world, like in a first person shooter (FPS) game or role playing game (RPG), moving the player's units in a real-time strategy (RTS) game, and moving objects in games where the screen does not move, like Tetris. Reasons as to why these actions warrant discussion is that certain storylines require the player to venture into a particular place before the event can happen; hence it would be beneficial for the player's gaming experience if that location is visited at a particular time during game play.

Resource Acquisition

Resources acquisition actions can be exemplified by picking up loot after slaying a monster in an RPG, picking up ammunition in an FPS, gathering resources in an RTS, consuming bonus items and collecting power-ups in most simple casual games, etc. There would be instances where the player is needed to pick up certain items that enhance the playing experience either by providing assistance to the player's current status or increase the player's sense of achievement when that particular item is collected.

Combat

This category consists of any kind of combat-related actions that include attacking or casting offensive spells on a monster in an RPG, shooting at enemies in an FPS, attacking another player's base in an RTS, etc. There are for example times when a particular entity has to be destroyed in order for the storyline to progress, and we need to get the player to perform these kinds of actions.

Accepting Requests

Among the most powerful type of actions that could be influenced is the acceptance of a request. While request acceptance is in most cases not the target action itself, following a request lead to the series of actions that results in the player experiencing various game content. Rather than trying to influence one action at a time, accepting requests can take players to the desired location, to pick up key items and to engage in combat with someone within the same context. Requests appear most often in RPGs as quests, and are fundamentally the same as missions commonly seen in FPSs and RTSs.

The next section covers the manipulation techniques, the implementations, concepts and some examples of possible implementations in games.

MANIPULATION TECHNIQUES

In this section, we propose a set of manipulation techniques. As mentioned above, these techniques are developed using psychological techniques found in other literature. The adaptation of the techniques is done by extracting the fundamental concepts that make the psychological techniques work, and modify them such that they will make sense in the gaming context. These concepts normally entail some requirements that must be fulfilled to be successfully applied, which also brings about difficulties in implementation in a gaming context.

A preliminary experiment was conducted using the techniques listed here to gain insights in the effectiveness of the adaptation of the techniques in influencing the player's actions. The techniques were implemented according to one of the described example implementations in a test scenario (a specially crafted Warcraft 3 scenario with a storyline). A total of 10 test subjects were chosen to participate in this experiment. They were chosen in a way to cover various demographics types, i.e., gaming habits of casual and hardcore, and different genders. Note that the results of the tests cannot be taken as a confirmation or rejection of the general techniques as they only exemplify one specific implementation. The tests also provide insights as to how the testing of the techniques can be implemented and be further improved.

For each technique, we will describe the general technique implementation, the concepts behind the proposed technique, implementation requirements, difficulties, and some examples of action specific implementations followed by the experimental results.

Commitment

For this technique, questions are to be asked to the player. The answers serve as a kind of commitment to the player, making the player responsible for holding true to his/her commitments, and thus perform actions that remain consistent with the commitment made. An example of such a question would be the player's views towards a subject, like liking of a particular non-player character (NPC) in the game. Then, if the player is presented with a kind of situation that is related to the commitment, we expect the player to choose or act in the way the player answered the commitment question before. The questions can for example be asked through a dialog or menu selections.

Concepts behind This Technique
People have the tendency to believe strongly in the things they have committed to, and they will remain consistent to what they have committed (Cialdini 2001). Making use of this mentality combined with other techniques like rational persuasion or getting the person lean towards a particular decision (Yulk and Falbe 1990), we can influence players in a video game to make a desired commitment for the commitment question and then have them adhere to this commitment with consistency.

Requirements and Difficulties with Implementation
The game must have some way of asking questions, like dialogs, and the player must have a means of making a decision, like pressing buttons for selecting different options.

Difficulties with the implementation arise with the additional dialogs to survey the player's preferences, which may disrupt game play, and that additional techniques may be needed to influence the player to choose the desired choice. A possible remedy for the second difficulty would be designing the game such that all the possible choices the player result in actions that are desirable.

Action-Specific Implementation Examples
Below are the example implementations for each action category this technique is capable of influencing.

Accepting Requests: To influence the player to accept quests given by a particular group of NPCs, the questions could be formulated in a way that would cause the player to commit to help this group of NPCs. Examples would be: "When asked for help by (a particular kind of NPC), would you help him or her?" or "Would you help the (a particular kind of NPC) when asked?"

Combat: To influence combat related actions, like attacking a type of NPCs unprovoked or avoid fighting with some kind of monster that the player is too weak to fight against, the question could be if the player would show hostility towards or avoid engaging certain groups of NPCs or the like. For example: “When faced with a stronger opponent, will you try to fight or to escape?”

Resource Acquisition: To influence the player to collect or gather a particular kind of resource that the player will need at a later point in the game, the question could be framed in a way that the particular resource is viewed to be in demand and then ask if the player will stockpile it. For example: “The price of (the resource) has risen considerably recently. Everyone is trying to get hold of it. Would you try to get your hands on (the resource) when you get the chance?”

Movement: To get the player to move to a certain place in the game using this technique, the question asked can be designed such that it mentions some characteristic of that location that would entice the player to go to that location. An example would be: “I hear that (the particular location) is a beautiful place. Wouldn’t you want to visit it?”

Results

Implementation Tested: Accepting Requests

Method of Testing: The player was asked by an NPC for his/her opinion on the types of critters that are present in the scenario – whether they like or dislike them. These critters can be found all over the game. The question serves to survey the player’s personal preference about seeing these critters. Thereafter, the NPC will ask if the player would partake in a quest to reduce the number of these critters due to a problem with overcrowding. The success of this implementation is determined by the pair of answers the player chooses; a success would be the like/refuse to kill and the dislike/kill pair of answers.

Sample mean, $\bar{X}$: 0.556

Probability of success, μ at 90% confidence level:

$$0.229 < \mu < 0.882$$

Discussion

From the empirical test results, the tested implementation of this technique does not display the effectiveness in manipulating players nor shows any capabilities of influencing the player’s follow-up attitude of the task, which deviates from the initial expectation. Reasons for this effect might be that the critters are merely decorative and do not affect the player, like attacking, giving experience or loot when killed etc. The player may not have developed any feelings or even notice the presence of these critters to warrant a strong reaction towards them. For the experimental test, the second question would however then need to be decoupled from these feelings to provide a valid test.

Return Favor

Have a non-player character (NPC) provide unconditional aid to the player in times of need. Later, when encountering this NPC again, the player should feel grateful towards the NPC for the help and should readily accept requests made by that NPC. The type of aid should be very useful in that particular situation, such as a full hit point (HP) recovery when the HP level is critical, and also might not even be material, like advice to get the player through tough times.

Concepts behind This Technique

People tend to express gratitude when receiving things from others. This concept of gratitude is internalized within most of us, and we actively seek to avoid labels like “ingrates” (Cialdini 2001). Exploiting this mentality to reciprocate, we can influence players to do things they would not do otherwise.

Requirements and Difficulties with Implementation

The game must be able to support NPCs, and the gratitude generated, which is the key factor of influence, has to be targeted at a specific entity or a group of entities. Without such entities, it will be very hard to invoke such feelings. In addition, as the action category being influenced is request acceptance, there is a need to support dialog as well as decision making in the game.

A difficulty for an implementation is the required additional scenes for the aid-providing event, adding more workload and time for the game developers for designing the scenes, debugging, etc. Measures must also be in place to ensure that the player triggers the aid-providing event, perhaps through the use of other manipulation techniques.

Action-Specific Implementation Examples

Below is an example implementation for the action category that this technique is capable of influencing.

Accepting Requests: To influence the player into accepting requests, one can implement an event where the player receives help by a specific NPC. For best results, one should provide help when it is critically needed to increase the sense of gratitude. Subsequently, one can present the player with a request by that same NPC that helped the player.

Results

Implementation Tested: Accepting Requests

Method of Testing: The player was placed in a situation where the player’s character is being chased by enemies. An NPC helps the player by providing a hiding place for the player and his companion, and leads the player’s enemies down a wrong path. After speaking to the NPC again, the NPC will have an option labeled “Return the Favor” which begins a quest, signifying the player is returning the favor by helping the NPC. The success of this implementation is determined by observing if the player chooses to accept the quest to return the favor.

Sample mean, $\bar{X}$: 0.7

Probability of success, μ at 90% confidence level:

$$0.420 < \mu < 0.980$$

Discussion

From the experimental test results, the tested implementation of this technique does not hint to the effectiveness in terms of player manipulation. Although the sample mean is high, the confidence interval has a lower limit of 0.420, which means that there is a chance that the actual probability of success is below 50%. Based on the results of a control experiment, the player may have chosen to accept the quest regardless of the aid. However, due to the small number of test subjects that triggered the control test, this inference may have no significance when tested with a large population of test subjects.

Time Pressure

Certain events in the game, where we want the player to participate, are only to occur for a certain interval of time, after which that same event has a lesser or no chance to occur again. The player should thereby be tempted to participate in this event because he/she cannot do so later (unless perhaps some game time has passed). Influence can be achieved on the player's actions at specific times, and as an additional benefit, the game developers can thus predict the times when those actions are executed and incorporate this information into their planning of the game.

Concepts behind This Technique

This manipulation technique plays with people's sense of value of a particular object due to its scarcity. In this case, there is a time constraint placed on the ability to participate in a certain event. The player should value the opportunity to participate in the event higher than what he/she is currently doing due to its scarcity (Cialdini 2001), and hence will feel pressurized (Yulk and Falbe 1990) to participate in the event before the time limit expires.

Requirements and Difficulties with Implementation

For a game that is going to apply this technique, the information that the time-limited event is going to take place must be conveyed to the player for this technique to be successful. Additionally, it can be useful to have an active timer being displayed to let the player know the remaining time.

Difficulties arise as the game must be able to limit access to certain aspects of the game, so that the relative scarcity can be achieved. The events also must be planned ahead of time. Triggering can either be achieved automatically after a certain time has passed or after some requirements are met.

Action Specific Implementation Examples

Below are the example implementations for the action categories that this technique is capable of influencing.

Accepting Requests: To influence the player to accept requests, the time-limited event could be implemented as the quest giver either only appearing for a limited amount of time or such that the quest is given out only during a specific time interval. For example, the player may be informed that during an eclipse, a particular NPC will appear and that talking to this NPC will result in a quest; or that a quest will be given by a particular NPC during the eclipse.

Combat: To influence the player into fighting particular monsters or a hostile NPC or a group of NPCs at a given time, the time that the monsters may appear can be constrained to a one-off or repeating time interval, for example, monsters appearing only at night. The player, who only has a limited opportunity of engaging the monsters, would be tempted to seek them out during this given interval.

Resource Acquisition: The resource that the player is supposed to collect can be made to appear with a certain timing, after which the resource will no longer be available or the player will have to wait for the next window of opportunity.

Movement: To get the player to visit certain locations at a particular time, access to those areas in the game can be blocked off and reopened during a specific time interval. The player would then be notified of the opening of the location and will be tempted to visit it before the location is closed off again.

Results

Implementation Tested: Resource Acquisition
Method of Testing: Upon triggering the event, an NPC will appear, telling the player that a time-limited event is going to occur. The player has two minutes to go to the designated area and collect gold coins that will spawn for the duration. Success of this implementation is determined by whether the player heads to the designated area before the time is up.
Sample mean, $\bar{X}$: 0.714
Probability of success, μ at 90% confidence level:

$$0.356 < \mu < 1.072$$

Discussion

From the empirical test results, the tested implementation of this technique does not seem to be effective to manipulate players. Although the sample mean is high, the confidence interval has a lower limit of 0.356, which means that there is a chance that the actual probability of success may be below 50%. The result of this test thus cannot deem the influence of the technique to be valid. The upper limit of the inequality is even (above) 1.0, which is a result of the high variance due to the small sample size. This effect also occurs in other calculations in this section. One shortcoming of our example implementation is certainly also that there are other means of acquiring the resource in the game, and that the player is aware of these other methods of getting gold.

Environmental Cueing

Environmental properties of the things that the player should pay attention to can be modified such that those objects or areas become more salient. This will generate more interest from players, and they are more likely to perform actions on these objects or to the area. For example, having sounds or voices played certain locations should capture the attention of the player and causes him/her to investigate its source. Another possibility is the use visual cues to stimulate interest in the area or an object. Through the prolonged use of this technique, the player will eventually get used to the effect and – with a coupling of corresponding rewards for the player – will then automatically execute the corresponding actions. This learning of cue-actions relations may be desirable depending on the context as the cues can be reused with increasing effectiveness.

Concepts behind This Technique

This technique relies on the fact that people perceive slight differences between two similar objects as being more different, thus more interesting and attention-grabbing, and ultimately influencing the player to become more aware of the surroundings and investigating the surroundings (Cialdini 2001).

Requirements and Difficulties with Implementation
The game must be able to present the necessary cues (visual or audio) to the player.

Difficulties in implementation include determining how salient the cues have to be to make them effective. If the cues are too salient, it might distract the player, while too subtle cues could be completely missed by the player.

Action-Specific Implementation Examples
Below are the example implementations for the action categories this technique is capable of influencing.

Combat: Certain hostile NPCs or monsters in a group can be changed to look or sound different, e.g., by adding an aura or some other minor details that sets them apart from the rest. Alternatively, unique music or sounds made by an NPC can be played when it appears.

Movement: In areas with hidden objects, NPCs or locations, special sounds or music can be played to attract the attention of the player and influence the player to investigate the area.

Results
Implementation Tested: Movement
Method of Testing: Upon entering a predefined region in the game, the player will hear a 3D sound being played. This sound is a human voice saying "Go away". Should the player move towards the direction of the sound, the player will find a hidden location with an NPC. Success of this implementation is determined by whether the player successfully locates the hidden area.
Sample mean, $\bar{X}$: 0.286
Probability of success, μ at 90% confidence level:

$$-0.072 < \mu < 0.644$$

Discussion
The empirical results of the tested implementation of this technique suggest that the test subjects were not affected by voices despite there being a dialog that hints the test subjects to attend to the voice. The sample mean is low for this implementation and the confidence interval has a lower limit of (less than) 0 which means that there is a good chance that the actual probability of success may be below 50%. The possible cause for this is that the cue is not salient enough to warrant interest in the player. Another thing to note is that the cue used in this implementation is an audio cue that conveys a message. The test is set up in a way such that the player is supposed to move to the area despite the message to prove that it is the cue and not the message that prompted the action. Visual cues, which are more salient than sound, may be employed with a greater success rate.

Group Influences

We can use multiple NPCs to influence the player to do a desired action by having these NPCs doing that same action in front of the player. Such demonstration events can for example be triggered when the player enters a new area, such that the player is not yet sure what to encounter.

Concepts behind This Technique
This technique makes use of the fact that people tend to follow what other people are doing in the event where they are not sure what is going on or unsure about how to approach a situation (Cialdini 2001). When the player sees many NPCs or other players doing a particular action, the player would likely also adopt the same behavior as the others.

Requirements and Difficulties with Implementation
NPCs that are to portray the action for the player need to be able to execute it, i.e., many more capabilities/animations may be needed for such NPCs.

Action-Specific Implementation Examples
Below are the example implementations for the action categories that this technique is capable of influencing.

Combat: To make the player feel the need to engage a particular enemy, one can have multiple NPCs surrounding a particular monster and attacking it.

Resource Acquisition: Having many NPCs using a particular kind of item, or letting the player see a scene of many NPCs collecting a particular resource, like gold mining, will influence the player to want to acquire that item/resource as well.

Movement: This implementation can be achieved by having many NPCs moving in a particular direction that leads toward the desired location that the player is to visit. The player should consequently want to move into that direction as well.

Results
Implementation Tested: Movement
Method of Testing: Upon walking into a predefined region, the player triggers a cinematic where three NPCs appear and begin walking towards a point that is outside the player's line of sight. The successful application of this technique is reported when the player moves towards the point the NPCs are moving to.
Sample mean, $\bar{X}$: 0.4
Probability of success, μ at 90% confidence level:

$$0.100 < \mu < 0.700$$

Discussion
The sample mean is very low for this implementation, and the confidence interval has a lower limit of 0.100, which means that there is a good chance that the actual probability of success may be below 50%. It can be inferred that the moving NPCs may not be an effective cue for this implementation and that the tested implementation does not

generate enough motivation in the player to follow the NPCs. There also may be a higher threshold of the number of NPCs required to influence the player.

Authority

By using NPCs that are perceived as leaders or experts for requests to the player, the player should be more agreeable to accept and follow through a task to the end. Examples of such leaders include scientists, squad leaders, a king, and basically any NPC who ranks higher than the player or NPCs who the player considers as superior (actual rank may not be important; qualities like experience etc. come into play).

Concepts behind This Technique
This technique is based off the fact that people tend to follow blindly to authoritative figures by assuming that the "experts" know what is best. There is also no need to prove the expertise of the expert. As long as the target perceives the person as an "expert", they will follow through the instructions without question (Cialdini 2001). The extent of this blind obedience to authority can be seen in the famous Milgram experiment (Milgram 1974).

Requirements and Difficulties with Implementation
The "experts" take the form of NPCs. The player must be able to converse with the NPC and choose whether to accept the request or not. Hence the game that wants to apply this technique must also cater for these functionalities. Alternatively, experts can also be used for giving pure advice, and leave it to the player to follow it or not.

A difficulty in implementation is the need to turn an NPC into a believable expert, i.e., the NPC must be designed in a way that clearly displays the qualities of an "expert". For example, by titles in the name, like "doctor", or the physical appearance of the NPC, like a leader that is easily distinguished from the other men. The mannerism of the NPC may also affect the effectiveness of this technique. Additional design considerations may thus be needed for introducing "experts".

Action-Specific Implementation Examples
Below is the example implementation for the action category that this technique is capable of influencing.

Accepting Requests: To influence the player into accepting requests made by a particular NPC, one can present the player with a request made by a NPC who fits the "expert" role. The request must be related to the perceived expertise of that NPC in order for the technique to work.

Results
Implementation Tested: Accepting Requests
Method of Testing: Upon triggering the event, a cut scene will play out, in which a group of six soldiers and their leader (distinctly different from the other soldiers) will march into the scene. The leader will stop and ask the test subject's character if he would join them to eliminate some monsters. Success of this technique is determined by the acceptance of the request made by the leader.

Sample mean, $\bar{X}$: 0.889
Probability of success, μ at 90% confidence level:

$$0.682 < \mu < 1.095$$

Discussion
This implemented technique fared better than many other techniques mentioned in this paper, as indicated by the high sample mean and the range of the probability of success. Based on the empirical results, it is inferred that players would react to requests given by a so called "leader" of the group of NPCs. Both appearance and mannerism (the way the leader speaks to his men and the player) were incorporated into the experiment.

Affective Musical Themes

This technique uses themed music to invoke feelings in the player to get them to react in the desired way. Examples include music with war drums to make the player feel like taking part in battle, or romantic music to get the player to interact with his/her love interest. Most games make use of music for experience enhancement, but they normally do not use it for the purpose of influencing the player's actions.

Concepts behind This Technique
People's moods tend to be affected by the music played in the surroundings (North et al.1999). Through this technique, we can get the players to develop certain feelings when we want them to. When people are in a certain mood, they tend to act according to how the general population acts when they are in such a mood ("feeling rules"), and we can influence the player's actions accordingly.

Requirements and Difficulties with Implementation
Music is the key feature of this technique manipulation; the game that wishes to employ this technique must thus be able to support music playback.

Learning can be incorporated into this technique to make the player associate a theme music with certain emotions (see the section on environmental cueing). It may take some time before the technique can be used with efficiency, and it may be difficult to determine when enough learning has happened to generate a good response.

Action-Specific Implementation Examples
Below are the example implementations for the action categories that this technique is capable of influencing.

Combat: Musical themes that are related to combat can be played whenever combat takes place. At the desired point to influence the player to engage in combat, this music is played and the player should react by engaging in combat.

Movement: Fast or slow music could influence the player's movement speed when transiting through areas where it would be beneficial if the player slowed down or hasten his/her pace.

Results
Implementation Tested: Combat
Method of Testing: The player is first taught to associate a sound with combat by playing a sound whenever the player engages in combat. Subsequently, when the player ventures into a predefined region, the sound will be played despite there being no combat. The success application of this implementation is determined by whether the player attacks a nearby group of monsters who are neutral to the player.
Sample mean, $\bar{X}$: 0.571
Probability of success, μ at 90% confidence level:

$$0.179 < \mu < 0.963$$

Discussion
The empirical results indicate that the tested implementation may not produce enough motivation to create the desired actions. The sample mean is close to 0.5, which may indicate that the implementation may achieve random successes.

Partner Requests

The player's partners can be used to make requests, which makes it hard for the player to refuse because of his/her relationship with them. Such NPCs includes party members, side-kicks, pets and love interests. These characters may accompany the player through most parts of the game; hence the player should develop some form of attachment to them and we can make use of this feeling of attachment to influence the player.

Concepts behind This Technique
The basis for this technique is the exploitation of the player's relationship with an NPC, which is formed throughout the game to influence the player. The acceptance rate for requests made by significant others is usually very high (Cialdini 2001).

Requirements and Difficulties with Implementation
Games without much interaction with NPCs and partners are obviously no good candidates for this technique.

Difficulties arise because this technique requires the player to build a relationship with the NPC, which requires a lot of effort to create situations where the player bonds with the NPC. The NPC must also exhibit mannerism similar to an actual person to make the relationship feel real enough to generate success. As relationship building takes time, the technique cannot be used on the fly.

Action-Specific Implementation Examples
Below is an example implementation for the action category that this technique is capable of influencing.

Accepting Requests: One can design an NPC that follows the player around, providing company to the player and building a relationship over time. The NPC then makes a request that the player should accept.

Results
Implementation Tested: Accepting Requests
Method of Testing: The player's character will be followed by an NPC throughout the game, playing the role of the player's love interest. This NPC will attempt to build a relationship with the player's character by conversing with the player about the various things they come across. Upon reaching the predefined area, the NPC will give a quest to the player. The success of this implementation is determined by whether the player accepts this request.
Sample mean, $\bar{X}$: 1.0
Probability of success: N.A.

Discussion
Based on the experimental results (100% successes), this is the most successful manipulation implementation tested.

Summary

The table below shows a summary of the results:

Technique	μ	$\bar{X}$
Commitment	$0.229 < \mu < 0.882$	0.556
Return Favor	$0.420 < \mu < 0.980$	0.7
Time Pressure	$0.356 < \mu < 1.072$	0.714
Environmental Cueing	$-0.072 < \mu < 0.644$	0.286
Group Influences	$0.100 < \mu < 0.700$	0.4
Authority	$0.682 < \mu < 1.095$	0.889
Affective Musical Themes	$0.179 < \mu < 0.963$	0.571
Partner Requests	N.A.	1.0

Table 1. Summary of the results of the experiment. μ is the success rate of the population based on a confidence interval of 90%, and $\bar{X}$ is the estimated success rate based on the sampling of the population.

CONCLUSION

In this paper, we proposed several player manipulation techniques that can help game developers to solve the problem of manipulating the player to behave in a specific way. Although many techniques of the experimental study show poor results, there were a couple of techniques that showed potential in the tested implementation. The small sample size, limited control experiments, and little experience in this new area of research have surely contributed to many of the poor results in the study so far. We would also like to highlight that this is just the start of a new field of research, and hope that this paper has generated interest in this exciting new area.

ACKNOWLEDGEMENT

This work was supported by the Singapore National Research Foundation Interactive Digital Media R&D Program under research Grant NRF NRF2007IDM-IDM002-051, as well as by a joint grant of Singapore's Ministry of Education / Academic Research Fund and the National University of Singapore under research Grant RG-263-001-148.

REFERENCES

Cialdini, R. B. 2001. *Influence: Science and Practice*. 4th edition, Allyn and Bacon, Massachusetts, USA.

North, A. C.; Hargreaves, D .J.; and McKendrick, J. 1999. The influence of in-store music on wine selections. *Journal of Applied Psychology* 84(2): 271-276l.

Milgram, S. 1974. *Obedience to Authority, An experimental view*, Harper & Row, New York, USA.

Yulk, G., and Falbe, C. M. 1990. Influence tactics and objectives in upward, downward, and lateral influence attempts. *Journal of Applied Psychology* 75(2): 132-140.

BIOGRAPHIES

Zi Xu, Siew

Alexander Nareyek

Zi Xu, Siew is a student of National University of Singapore, majoring in Computer Engineering. The techniques proposed in this paper were developed as part of the final year project on the topic of Player Manipulation initiated by the NUS Games Lab.

Alexander Nareyek is the director of the interactive intelligence Labs (ii Labs) and of the multi-disciplinary Games Lab at the National University of Singapore, where he also holds an assistant professorship. He received his diploma and Ph.D. from the TU Berlin/Germany, held positions at GMD-FIRST/Germany, Carnegie Mellon University/USA, and the Cork Constraint Computation Centre/Ireland, and served as CEO and CTO for Digital Drama Studios/Czech Republic. He is the founder and owner of the games company Kingdom Crafting Pte Ltd, and also serves as chairman of the International Game Developers Association's Artificial Intelligence Special Interest Group.

The Dark Side of MMOGs: Why People Quit Playing

Don Anderson
Intellas Group, LLC, 9106 Bristol Avenue, Louisville, KY 40220 USA
(502) 499-1800
danderson@intellas.com

Keywords

Games, social, psychological, MMOG, interaction, community, Tribal Wars

Abstract

The attraction of MMOGs is obvious as a broader range of people are drawn to the interactive medium for a variety of reasons. This is in particular evident to me to see my 30 something daughter and my teen age grandson both involved in "World of Warcraft." I, the father and grandfather, participated actively in another MMOG, "Tribal Wars", for over a year. However there is a negative or dark side to MMOG involvement, and multiple reasons that players eventually leave individual games or the genre in general. The purpose of this paper is to examine some of these issues that the author has seen both firsthand and also in discussions with others who have been or are currently involved in the genre.

Introduction and Background

Interactive gaming has surpassed the movie medium in attracting new participants especially amongst the younger generations. As stated before in previous studies, one large MMOG, "World of Warcraft," has managed to make in one year subscriptions more revenue than the largest grossing film of all time, "Titanic." There is no doubt that the attraction of games and media that allow interaction in which the viewer is also a participant will continue to grow in popularity as at the same time the older more passive forms of entertainment will diminish in relevance and size of audience.

Despite this increase in audience and popularity there are still serious issues with the long-term viability of many of the existing MMOGs. One of the primary challenges is providing new experiences for users to keep them engaged in the particular world in which the MMOG is staged. A lack of new experiences can lead to stagnation in any game and subsequent loss of interest on the part of participants.

Some observations to these problems and the success or lack of success of several MMOGs to address them are the subject of this brief study. The author will cite both personal experience as well as the experience of other participants in the MMOG worlds.

Different Game Worlds – Different Problems and Issues

Different MMOG world scenarios present a variety of different issues and problems depending upon the nature of the game. Role Playing MMOG's such as World of Warcraft and Everquest differ in structure and interaction from Strategy games such as Tribal Wars. In a typical RPG a participant plays a single character at any one time, although many people involved in these RPG's will have several different characters of different races and character types. Running a single player character also involves forming alliances and joining groups such as guilds to coordinate and cooperate with other players in the game. This involves coordinating playing times for joint operations such as raids.

Strategic MMOGs such as Tribal Wars pits the player against other players as a tribal warlord or medieval lord. Operations are more strategic in scope whereas the RPG interactions tend to be more of a tactical nature. The player in the strategic game positions himself with other players as either a member of the same tribe, an allied tribe, a neutral tribe or an enemy tribe.

Although there are differences in the various types of MMOGs that players become involved in, there are some common threads and problems that span across all genres and impact the paradigm of MMOG gaming as a whole.

Common Issues and Reasons Why People Quit MMOGs

The author has identified four primary reasons that lead to participants leaving MMOGs. These are not all inclusive and others may be added or discovered in the future, but they do give some insight into the problems inherent in this form of gaming. Also, just because a player leaves a particular game world does not mean that he leaves MMOG gaming altogether. Sometimes a player leaves and then joins another MMOG world. The four major problems seem to be:

1. No end to the game or lack of clearly defined goals
2. The amount of real world time that the game demands for the player to be affective in the game world
3. Interference with the gamers' personal life
4. Overly complex and/or oversized game world

Cameron Sorden has another list of game exit points in his article on Massively.com and classifies the problems into three different categories:

1. Gameplay Issues
2. Social Issues
3. Personal Issues

This is simply another way to view the issues and problems in the MMOGs. The previous four would easily fit into this classification as well.

We will look at each of these four issues separately, although many times they are related to one another.

No End to the Game or Lack of Clearly Defined Goals

Any game that attracts a new subscriber and participant is usually interesting and engaging at the onset. The player has plenty of new territory to explore and new facets of the game to learn which leads to continued engagement and interest. This "honeymoon" period can last for weeks or even months, but as the player grows either his character or his empire within the game world, the initial fascination can quickly fade unless there are new challenges to meet.

This problem is particularly evident in strategic games such as “Tribal Wars.” Once a player has successfully survived to be a power in the game world in which he is playing, the focus can be muddled by such factors as problems within the tribe, other tribe members and friends leaving the game, the sheer size of the empire, and the complexity of management of game tasks, among others.

Successful MMOGs such as “World of Warcraft” avoid some of these issues or at least delay the problem by creating new modules which enhance the game by adding new areas to explore, new equipment to acquire, etc. Even this can still not assuage some players from getting burned out and leaving. The author recently encountered a seminary student who was working at the local Best Buy store. In our conversation our mutual interest in gaming led us to discuss the fact that he was a subscriber to “World of Warcraft.” The young man stated that he was getting ready to leave the game because “I acquired all the equipment that I could and can not advance any more levels.”
In my year and a half playing “Tribal Wars”, I personally experienced similar burn out and lost interest when a number of my friends and allies left the game. Managing my 350 villages became a chore each day and without the social interaction with the people I enjoyed playing with most, and with no clearly defined end game in sight, I quietly left the game.

The amount of real world time that the game demands

MMOGs typically require a significant amount of real world time in order to be successful in the game world. Also the social aspect of the game puts increasing pressure on the participant to “be there” when the tribe, guild or party wants to take action. Initially the immersion into the new game world can be exciting, but in time the demands of the real world can begin to clash with the time spent in the game.

My own experience in “Tribal Wars” clearly demonstrated to me this problem inherent in most of the MMOGs. I found that I had to check my villages and game status from the time I got up in the morning until I went to bed at night. Several issues in the game even had me waking up in the middle of the night to perform a procedure, such as dodging or launching an attack. As the size of my holdings in the game world increased, this became more time consuming. I soon found that I was spending hours each day in the game and it was interfering with my life in the real world.

Eventually I had to make a decision to leave the game altogether because the game tasks were becoming both tedious and way to time consuming.
On visiting with my family on a recent vacation trip I observed my teen age grandson staying up to the wee hours of the morning engaged in “World of Warcraft.” His mother originally introduced him to the game as something they could do together, but now when he is at home, she rarely can get online to play her game. The addictive nature of these games lends itself to these problems.

Interference with the gamer's personal life

As noted in the section on the problematic amount of time that most of these MMOGs demand, leads directly into the next problem. In discussions with tribe mates in "Tribal Wars" some of the related problems became apparent. One young man who was a long time friend and tribe mate was very active in the game for many months, but was engaged to be married. Immediately before the wedding, his time in the game became more sporadic and after his wedding, he finally dropped out of the game due to the demands of married life. Other men who had been married for sometime were being pressured by their wives to spend less time in the game.

If both marriage partners are interested in the game and are both engaged, this can eliminate this particular problem. I have observed a few cases in which husband and wife both play. This is probably more common in an RPG like "World of Warcraft."

In "Tribal Wars" a couple of tribe mates left the game to care for sick relatives. So once again the demands of the real world can impact the will or ability of a player to continue in the game world in which they are involved. Different issues and demands come into play based many times by the stage of life of the gamer. The phenomenon is similar to other activities and hobbies in life. As a child I began to collect postage stamps and continued until I discovered girls as a teen. After marriage I resumed the hobby until the children came along and most of the time and money was spent with them. After a divorce and the kids getting older, I resumed the hobby for a time again. Similar cycles appear to also affect gaming.

Overly complex and/or oversized game world

This problem became apparent in "Tribal Wars" the longer that the author was involved in the game. Each world is so big that either you become isolated from action in a particular section of the world or you are so spread out that management becomes a major chore. Part of the attraction of smaller scale games with limited size and scope is that the tasks are more manageable and the goals are more clearly defined. In many MMOGs the goal is to either increase your holdings and number of villages as in "Tribal Wars" or to increase your level, possessions and personal influence as in role-playing MMOGs such as "World of Warcraft."

The average player seems to reach a point where enough is enough, and has to make a decision when to leave the game. Some players solve this by starting accounts in other worlds of the same MMOG or creating other characters to develop. One solution doesn't fit all, but change of some kind is usually inevitable.

The other issue that the author has observed is the complexity of learning game play. The one game world that I found this to be a particular issue in was "Entropia Universe." The first issue is the large complex software that this MMOG requires on the local system. The next issue is the focus on monetization within the game world that either requires the player to deposit sums of cash to buy game currency or to labor harvesting "sweat" from strange animals in the game, a process that was hard to figure out exactly what to do. The complexity was a game breaker for me.

Conclusions

Despite the problems inherent in online MMOG gaming, this game style will continue to grow and dominate well into the future. The opportunity for the creation of even better and more full featured games is virtually endless and will continue to attract a following based on the interest generated by the new game. Gamers are always looking for a new challenge, and a more immersive experience.

The problems inherent in these games will persist and will challenge the crafters of new game worlds to address these issues. People will continue to play in a dynamic environment that provides new and interesting experiences, and will grow bored and seek new outlets when the game becomes stale and static.

It is obviously far easier to retain an existing customer than to recruit a new one. MMOGs such as "World of Warcraft" that pay attention to the reason that subscribers leave the game and seek to address the issues that are in their control will continue to succeed in the genre, whereas MMOGs who fail in this mission will fade away and be replaced by newer and better games.

References

D. Anderson, "Social and Psychological Dynamics in MMOs," Proceedings of CGAMES 2008, July 30 – August 2, 2008, Louisville, KY, USA

N. Ducheneaut, N.Yee, E.Nickell, R. Moore, "'Alone Together?' Exploring the Social Dynamics of Massively Multiplayer Online Games," Proceedings of CHI 2006, April 22-27, 2006, Montreal, Quebec, Canada

N. Ducheneaut, R. Moore, "The social side of gaming: a study of interaction patterns in a massively multiplayer online game," Proceedings of CSCW 2004, ACM, New York, pp.360 – 369.

M. Jakobson, T. Taylor, "The Sopranos Meets EverQuest: social networking in massively multiplayer online games," Proceedings of DAC 2003, Melbourne, Australia, pp.81 – 90.

N. Lazzaro, "Why we play games: four keys to emotion without story," 2004

C. Sorden, "Player vs. Everything: The retention game," Massively.com, June, 2008

D. Warner, M. Raiter, "Social Context in Massively-Multiplayer Online Games (MMOGs): Ethical Questions in Shared Space," International Review of Information Ethics, Vol.4 , December, 2005, pp. 46 - 52

N. Yee, "Ariadne – Understanding MMORPG Addiction," http://www.Nickyee.com/facets/home.html

CREATING A DESIGN FRAMEWORK FOR EDUCATIONAL LANGUAGE GAMES UTILIZING HYBRID MEDIA

Janne Paavilainen[1], Hannamari Saarenpää[1], Anu Seisto[2] and Maija Federley[2]

[1]University of Tampere
{janne.paavilainen, hannamari.saarenpaa}@uta.fi
Department of Information Studies and Interactive Media, Games Research Lab
FI-33014 University of Tampere
FINLAND

[2]KCL
{anu.seisto, maija.federley}@kcl.fi
Oy Keskuslaboratorio - Centrallaboratorium Ab
P.O. Box 70, FI-02151, Espoo
FINLAND

KEYWORDS
Video Games, Educational Games, Language Learning, Hybrid Media, Game Design, Design Framework.

ABSTRACT

Educational games have seen grown interest in recent years. At the same time, there has been interest in the research of hybrid media. Hybrid media is a combination of print and electronic media. In this paper we present a first study on the design of educational language games for hybrid media platform. Based on scenario interviews and supportive research, we present a design framework which can be used to aid the design of educational language games. For our study, we interviewed six elementary school teachers and 16 6th grade students aging from 11 to 12 years. We also interviewed four study book authors and examined educational language games.

INTRODUCTION

The use of video games in education has seen growing interest in recent years. The importance and capabilities of educational video games have been noted by many authors and games are considered to be efficient tools for learning, as they are able to motivate the learner (e.g. Prensky 2001, Kiili 2005, Gee 2007, Annetta 2008). Aguilera and Méndiz reviewed the research on the subject from 1970's to 2003 and concluded "[F]or learning, video games are of unquestionable importance, and can be used… at different academic levels." (Aguilera and Méndiz 2003).

Hybrid media, i.e. the combination of print and electronic media, has been a popular research field in Finland and around the world. Especially in Finland, the combination of print products and mobile technologies has gained great interest, as these two industries are important for the national economy (Seisto et al. 2009).

The combination of educational games and hybrid media has not been studied rigoursly before. Game based learning and new technology solutions could open up new possibilities in education. Educational games combined with hybrid media solutions is an interesting research area as educational games are gaining more interest and print media has still an important role in schools. Earlier study showed that teachers are interested in using hybrid media in education (Seisto et al. 2009).

Learning by Hybrid Media (LehMa) is a project where the aim is to develop educational games for a hybrid media platform. The current research focus is in English education for 6th grade elementary school students (11 to 12 year olds). This paper presents the design framework for educational language games, which was created in the early phase of the project in the end of 2008. The design framework was later applied in the development of educational language games for hybrid media platform. In this paper we focus on describing the creation of the design framework.

The design framework was created based on scenario and supportive research. The scenario research consisted of interviews with six elementary school teachers and 16 pupils. The supportive research was based on literature review, interviewing the authors of the English study book currently used in many elementary schools in Finland and carrying out an informal examination on the current educational English learning games available from a Finnish online education service.

HYBRID MEDIA

There are many concepts and words to describe the combination of two or more media (multichannel, multiplatform, cross-media, integrated media etc.). Most of these definitions exclude print media however. Concept of hybrid media is used to describe the combination of fiber-based and other (electronic) media. But it is not unambiguous definition either. Different levels of convergence between print and digital are included under single term of hybrid media. So, there are different levels of hybrid media (Figure 1 on next page). A common aspect for these different levels is, however, that hybrid media means an application that combines the printed and electronic media in a way that the service is complete and adds value to end user only when both media are used.

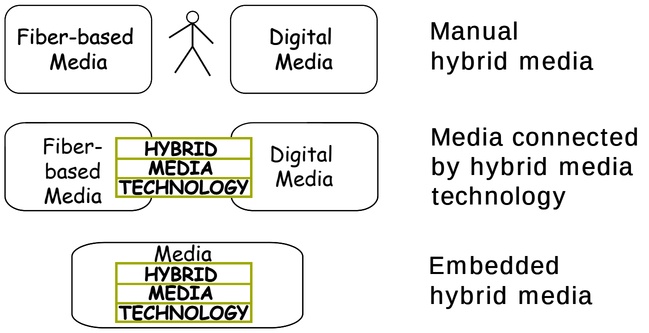


Figure 1. Different levels of hybrid media concept (Oittinen 2006).

RELATED WORK

Games have a lot of potential to enhance learning. This is because they have the ability to motivate and engage people (Bogost 2007). People have intrinsic curiosity (Malone 1980) which makes them want to try and master challenges. Games provide these challenges in a meaningful environment which also allow people to try out things that would not be possible otherwise. For example in a game a player can experience historical events that happened long time ago.

Educational games have been studied a lot and their potential to assist learning has been recognized widely. Still the design of an effective educational game remains difficult. The main reason why there are not that many good educational games available is the gap between educators and game designers. Educators usually know how to teach efficiently but they do not understand game design and vice versa. This has to change in order to improve the quality of educational games (Fortugno and Zimmerman 2005).

One of the first design heuristics for games were presented by Malone in 1980. According to Malone, the three essential design principles for good games are *challenge*, *fantasy*, and *curiosity*, and each of them contains several sub-components (Malone 1980). After Malone's design heuristics, many other guidelines for both evaluation and design of leisure and educational games have been presented (e.g. Federoff 2002, Kämäräinen 2003, Desurvire et al. 2005, Kiili 2005, Korhonen and Koivisto 2006). Evaluation guidelines are mentioned here because they can be also used in design.

Prensky has studied digital game-based learning and identifies several factors on what makes games engaging and what contributes as good game design (Prensky 2001). Gee has done a bit similar work and presents learning principles that good games incorporate (Gee 2003-2007). He presents 16 principles which are characteristics for good video games and these should be also found from educational games. Both Prensky's and Gee's listings give aid for educational game design, but they are not very concrete or practical from game design point of view.

Henriksen presents eight dimensions for educational games, which can be used in the development process (Henriksen 2006). These dimensions have been created to bring education and games closer each other. Henriksen's design dimensions are presented in the Table 1 below.

Table 1. Henriksen's eight dimensions for educational game design.

Content – Participation	Balancing the focus on content and participational incentives as well as optimizing the relationship between the two.
Narratology – Ludology	Considering the approach of games, distinguishing between the story told and the rules manipulated.
Endogenous – Exogenous incentive	Considering the level of integration between participational incetives and content providing elements.
Construction based - Railroading	Considering the degrees of freedom of the player, as well as the origin of legitimate knowledge in the learning game.
Analogue - Digital	Considering the media utilised for the game, as well as crossmedia considerations.
In-game – Off-game learning	Considering the placement of the learning effort compared to the game element.
Simplicity – Complexity	Considering the balance between representational complexity and ease of use.
Reflection - Flow	Considering the desirable state of mind while playing the learning game.

Henriksen's guidelines do not try to give exact instructions on how to design a learning game but to offer assistance and directions on what to focus during the design process.

Thomas has designed guidelines for designing pervasive learning games, which means educational games that are on around the clock and have context sensitive features (Thomas 2007). She has identified four categories which include total of thirteen principles that should be followed when designing pervasive learning games. The categories are: *community*, *autonomy*, *locationality*, and *relationality*. These principles are intented specifically for pervasive games but most of them can be used also in other types of educational games. The community category includes three sub-principles and highlights the importance of collaboration between players and that the game should support that. The second category is autonomy which highlights freedom, control and personal growth. The third category, locationality is about spatial and temporal freedom of a player and the use of context. And finally, the fourth category, relationality, stresses personal environments, goals, and meanings. These thirteen principles are already quite precise and they offer a good assistance for design of educational pervasive game. Still these guidelines are too specific to be used fully in other educational contexts.

Although there are various frameworks, heuristics and guidelines available, we decided to create a new design framework for several reasons. First, none of the earlier guidelines focus on language learning games. Second, the practicality of the earlier guidelines is questionable when considering the actual design process and concrete solutions to design questions. Thirdly, none of the guidelines consider hybrid media in any form. Therefore we decided to build our own design framework, which would be based on user-centred design.

RESEARCH METHODS

The basis of our research was user-centred design as we had no prior experience in creating a design framework for educational language games. Rather than relying solely on our own preferences and literature, we saw the end-users (i.e. teachers and pupils) as invaluable informants.

User-centred design and scenario research were succesfully used in an earlier research project in the *Games Research Lab* (Ermi and Mäyrä 2005). Like in *LehMa*, the earlier project featured the use of new technology and game concepts. This encouraged us to follow similar path and selecting scenario research as our main research method. We did also supportive research in the form of literature review, interviewing the authors of an English study book and informally examining educational language games for learning English.

Scenario Research

According to Ermi and Mäyrä (2005) scenarios are prevalent way to bridge gap between users and designers.

> “They are short fictional narratives that describe a use situation and the interaction between users and and proposed systems, and can be used to discuss and picture different kinds of future-use situations of technology” (Ermi and Mäyrä 2005).

Like Ermi and Mäyrä, we found it important that the concrete nature of scenarios help to create a common understanding of the proposed technology and it’s capabilities. Through scenarios, it would be easier to discuss about educational games and hybrid media with the interviewees. The goal of the scenario research was to gain ideas and inspiration considering the design framework for educational language games.

We had both adults and children as our informants, so we decided also to use comic-strip scenarios, which are illustrative and easy to comprehend like stated by Ermi and Mäyrä (2005). Although the scenarios portraided finished game concepts and complete hybrid media solutions, their purpose was to act as starting points for the interview discussions. The discussions were considered as the most important data in the study as they were used as the basis when creating the design framework.

Designing Scenarios

The comic-strip scenarios were created inside the research team in collaborative brainstorming sessions. There were numerous scenario ideas, but eventually they were numbered down to six different scenarios. During the supportive research, early versions from the scenarios were shown to the book authors, whose comments were used to finish the scenarios into their final forms. The scenarios featured different kinds of use situations both in school and at home, which featured educational games, language learning and hybrid media in different forms. Every scenario did not combine each of these three elements and some of the scenarios were more provocative than the others.

During our research there were no technological specifications available for the hybrid media platform, which would be used later in the project. Therefore we were unable to attach any specific technological solution to the scenarios or to the design framework. However, for the scenarios, we defined three different hybrid media fidelity levels, which would represent the overall features of different hybrid media solutions. The fidelity levels were named as low, medium and high accordingly.

The low fidelity level is based on using two-dimensional codes in the pages of a printed study book. These two-dimensional codes can be read with a mobile phone, which has a camera and identification software. The user would focus the mobile phone’s camera on the two-dimensional code, which would then start an educational application which runs in the mobile phone. The earlier study which probed the teachers’ attitudes towards hybrid media was based on this fidelity level (Seisto et al. 2009).

The medium fidelity level is based on printed study book which contains electronics. The electronics would be different kind of radio transmitters, using Bluetooth technology for example. Instead of two-dimensional codes in the study book, the pages would feature a touch-access button which would activate the radio transmitter. The radio signal would then start an educational application in a mobile device or in a computer.

The high fidelity level represents a blue sky solution where the study book would be a standalone platform, losing some of its “hybrid” features as everything would be packed into a one device. The high fidelity level represents a thin bendable full color display with touch sensitive interface used by either hand or stylus pen. This bendable display would also contain multimedia capabilities with audio input and output. The high fidelity level is a vision from a futuristic study book, which would be available in the years to come. Although this fidelity level unreachable in the scope of *LehMa* project, it was illustrated in the scenarios to inspire discussion and widen the design space.

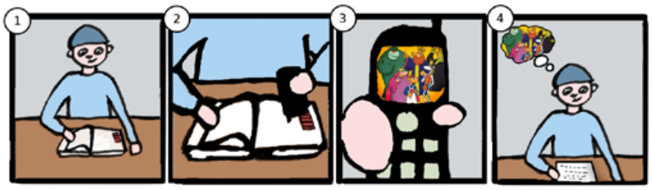


Figure 2. An example scenario featuring low fidelity hybrid media solution.

It is not possible to go though all the designed scenarios in the scope of this paper, but one example is shown above in Figure 2. In this scenario, the pupil scans the two dimensional code found on the study book's page. As a result, a video is shown which illustrates an important grammar issue in English language. The last frame shows that the pupil is doing good in the following day's exam as he remembers the exciting learning situation with hybrid media. In our scenario illusstrations, there were descriptive text beneath each frame explaining the scenario in narrative manner.

User Study

The user study consisted from scenario interviews, where each scenario was shown to the interviewees and discussed accordingly. We interviewed six elementary school teachers (three male, three female) and 16 elementary school pupils (eight boys, eight girls). The pupils were in the 6th grade (11-12 year olds) and they all studied English as their second language.

The teachers were interviewed one-on-one, but the pupils were interviewed in pairs by one interviewer. It has been noted in other studies, that interviewing young children in pairs makes them more relaxed and open for discussion (Höysniemi et al. 2004).

During the interviews each scenario was shown to the interviewees one by one. After examining the scenario, the interviewee filled up a short feedback form and the interviewer continued with specific questions relating to the scenario at hand. This was repeated until all six scenarios had been shown. The interviewees ranked the scenarios from best to worst and after the scenarios, the interviewer asked general questions relating to games, education and hybrid media. The interviews lasted from 35 to 50 minutes and they were recorded for further analysis.

Supportive Research

The supportive research consisted from three different smaller studies. The first was the literature review on educational game design, which was discussed in the related work section earlier in this paper. The second was the interview of the book authors of the current English study book used in many elementary schools in Finland. The third was the informal examination of current educational games used for teaching English in Finnish elementary schools.

Authors Interview

The research team interviewed the four authors of the current English study book, which is commonly, but not exclusively, used in Finnish elementary schools. The purpose of the interview was to discuss about the early versions of the scenarios and the use of educational games in learning English in general.

Game Examinations

The informal examination on educational English learning games was performed to achieve an understanding from the current situation. Due the short timeframe in the early phase of the project, it was not possible to make a complete formal analysis based on several different sources. The research team selected one major Finnish online service, which provides educational games for many subjects including math, biology, physics, foreign languages etc. This particular service was selected because it is well known and it is used by the school which teachers and pupils were interviewed in the scenario research. The researchers, who have expertise in game design and evaluation, focused on several games which were related to studying English and played them until they had clear understanding from them.

RESULTS

In this section we present the key findings from the scenario and the supportive research.

Scenario Research Results

From the 84 instances where a scenario was presented, there were only six cases where the interviewer had to clarify the scenario for the interviewees. In addition, only in three cases the interviewees asked a question related to the shown scenario. This implies that the scenarios were well understood. The comic-strip format proved to be useful and both the teachers and the pupils were able to give feedback based on the scenarios.

All the interviewees had positive attitudes towards using video games in education. Regardless of gender, hybrid media was seen as an interesting platform, especially among the pupils and in the high fidelity format. The teachers and the girls stated that the new technology and video games intrique especially boys. Similar gender based statements from the teachers and the girls were apparent through-out the interviews, as many times it was considered that the boys would be the most excited, or that they would get the most benefit from implementing new technology and video games into education.

Some teachers and girls stated that the new technology should be actually useful in a new way, meaning that the technology actually brings something new to the learning process. One teacher had an idea of improving boys' handwriting skills with a stylus pen and high fidelity hybrid media technology. Some teachers considered that the presence of technology might spring up disruptive behaviour. For example, the pupils have the tendency to start playing around with their mobile phones in a class room (hence their use during a class is widely prohibited in Finnish schools). This disruption might be even worse problem in a home work situation. Also, when working with mobile phones, there is the challenge with various models, which might result into unequal learning opportunities.

Currently the school under our study allows the use of mobile phones in some classes. In math for example, mobile phones are used as calculators. Sometimes the pupils use their mobile phone cameras to photograph their homework from the chalkboard. Although new technology has a novelty value, some pupils noted that educational games should also have depth, because the novelty value will run out soon. Teachers also considered that innovative technology has a positive effect on motivation to learn, but it might not last for long.

Playing video games in general was considered to be good for learning English. Not surprisingly the boys were suggesting that it would be fun to play shooter games during classes. However, they were also mature in their answers in a sense that they also considered educational aspects in games. One of the boy pairs stated that multiplayer shooter games teach you squad tactics, communication and collaboration. Some students considered that as adults do not generally play games, they do not understand them.

Current educational games were considered boring by both teachers and pupils. They were considered too simple and repetitive. One teacher expressed that there is a too large cap between commercial leisure games and educational games. Some of the students mentioned this also, but added that playing inferior educational games in school is always better than no gaming at all. Both the teachers and the students stated that usually the online service is used once in a week for the time of one class i.e. 45 minutes. Some teachers and students felt that this was too short time. When discussing the length of a one play session, 15 minutes was considered appropriate. This means that after 15 minutes, the game is changed so that the interest lasts.

Everyone wanted more complex games. Especially open-ended virtual worlds or world building games were considered to be good in education. Games which have persistent worlds and continuity were favored over simple games that one must start from the beginning in every class. Commitment and long-term character development was praised also. Interestingly it was considered that virtual worlds and characters could be used to teach some of the more difficult topics, like ethics and morality, in practice. One statement from an English teacher was that educational games for English should try to foster the three-layer model in language learning: identify, apply and produce. She added that educational games should make the pupil to think and not mechanically perform tasks. However, mini-games, i.e. short games which are quick to start and stop, were seen useful especially in a group work situation where one group is finished and waiting for the others to finish. In this situation, educative mini-games would be a good way to pass time and keep the pupils occupied, which would then maintain order in the class also. From the classic study book tasks, the students stated that crossword puzzles and other vocabulary tasks, such as word circling, were the most enjoyable to do.

Some teachers considered that educational games could be used as a substitute for formal exams when applicable. The teachers, and some of the pupils, saw that one of the benefits of educational games would be their ability to give instant feedback and automate error checking, which is usually very laborious task for the teachers. Automatic differentiation, or adaptive difficulty, was seen as important feature in educational games by the interviewees. Another important feature would be the possibility to monitor, evaluate compare and report the performance of the pupil. One teacher stated that educational games in general should be simple enough, so that the older generation of teachers could understand them. This challenge between the generations will most probably fix itself in the future as there will be more teachers with higher media and video game literacy.

Collaboration and group work was emphasized by the teachers and the pupils. Sometimes it was considered to be a bit laborious, especially for the teachers. Both the teachers and the pupils promoted peer support and one of the boy pairs stated that you learn yourself when you are helping others. Generally, social interaction in class work was valued highly by everyone. The girls were as excited about group work as the boys, but many girls stated that boys tend to disrupt the group work process in the class. Some focus orientated tasks, like listening comprehensions, were stated to foster order in the class. Some teachers noted that listening comprehensions, which are held in a normal class room instead of spesifically designed audio class room, might be awkward to organize if every pupil needs to have their own gear for it.

Extrinsic motivational factors were considered good and exciting. Especially if it would be possible to mix educational gaming at school and leisure gaming at home. In this example, playing an educational game in school would also benefit leisure gaming at home. One girl pair stated that if you could get points from the school and be able to spend them at home at your favorite game, it would be a good method to show to your parents that you are doing well in school. One teacher considered that this might be too much gaming though. The pupils considered that even boring subjects would be a lot more fun if there are special rewards present. However, these rewards should be possible for everyone to achieve, so that that the class would not split into elite pupils and less fortunate ones. One female teacher noted that there are clear motivational differences between boys and girls, and the rewards should be considered accordingly. One boy pair stated that focusing too much on the reward might actually be counter-productive, as you would not necessarely learn anything then. One of the teachers stated that her class collects “smiley faces” which are achieved by doing your home work. Certain amount of “smiley faces” results into a reward and the pupils seem to be motivated by this practice which is basically a long-term reward mechanic.

Supportive Research Results

In this section we present the key findings from the authors' interview and from the informal examination of current educational games for learning English.

Authors Interview Results

During the book author interview, the authors highlighted three issues concerning educational language games. The first issue related to the use context of educational English learning games. During classes, the emphasis should be in social interaction and in the verbal use of the language. This also emphasizes hearing comprehension tasks. Homework is naturally focused more on reading and writing. The second issue considered differentiation, which was seen as very important and the authors told that traditionally there has not been enough differentiation for the advanced students. This could be seen in practice also as the authors' study book features extra materials for the slower learners but not for the faster ones. The third issue brought up by the authors related to the design of educational games. They should be simple enough so that the older generation of teachers are able to understand how they are played and what they try to achieve from the perspective of learning. The research team showed the early versions from the comic-strip scenarios and the authors considered that the they were understandable and effective for creating feedback and discussion.

Game Examinations Results

Examining the educational English learning games brought up several issues which should be taken into consideration when designing educational games.

The examined games resembled classical titles such as *Pac-Man* (maze), *Super Mario* (platformer), *Mah-Jong* (puzzle) and *Tic-Tac-Toe* (parlour). The most basic issue was that the gameplay and learning were not intervowen in any way. For example, in a Pac-Man style maze game (Figure 3) the player guides his character through the maze, collecting valuables and evading the ghosts. When the player pics up a certain item, a dialog pops up with a translation exercise, where the player is asked to click the corrent word from three possibilities. Regardless of the outcome, the player continues the maze game after choosing the word. In this case, the game play and learning are separated and they do not seem to support each other in any way. The player does not really learn anything while guiding the character in the maze and the pop-up dialog interrupts the gameplay.

Figure 3. Maze game with a translation exercise.

The online service keeps a record on when and how many times the player has played the games Otherwise there were no support for continuity or persistant game play. If the player comes back for a game, she has to start all over again. In addition to the repetitive nature of the games, many of the games featured the same game mechanics and the only difference was the visual outlook. Some of the games were really short and it took more time to getting the games started than actually playing them. There were no differentiation used in the games in the form of adaptive difficulty levels. Lastly the games suffered from basic usability and playability problems. For example, sometimes the games ended up in stalemate, and the only option for the player was to reset the game and start all over.

DESIGN FRAMEWORK

The findings from the scenario and from the supportive research were transformed into a design framework containing 10 items (see Table 2 on next page). The findings were used as inspiration for creating the framework and they were not considered as be-all-end-all laws of design. Naturally, the research team's expertise in game design and evaluation is shown in the framework and it was created to cater the needs of *LehMa* project. However, it was also our goal to produce practical and illustrative design framework which could be used by others as well. Notice that the the framework is not tied into any technological solution. This was partly because at the time of the study, there was no certainty on the technology solution which would be used later in the project. We also wanted to keep the design framework free from any technological constraits. This design framework table consists from a running item number, item title, short description from the item and a practical example on how the item could be used.

DISCUSSION

In this paper we have presented a design framework, which has been and can be applied in the development of educational language games. The role of hybrid media could be seen as irrelevant in this paper, especially as none of the framework items are directly related to the technology platform. However, as the ultimate goal is to create viable educational games for hybrid media platform, it would have been unwise to exclude the technology from the scenarios for example. We did not have the knowledge how it would have affected on the interview results nor we did not know at the time what would be the actual technology in the development phase. Therefore, it was a safe bet to implement hybrid media in the scenario interviews, although the technology or its possible capabilities are not presented in the framework.

In regards to hybrid media, our findings are very general and somewhat vague. All teachers and pupils seemed to like the idea of using new technology in learning. The ease of use, ability to network with both in school and at home, technological equality among pupils etc. are all expected and axiomatic. When the actual hybrid media platform is ready and available, only then it is possible to evaluate it and create design guidelines aimed at that specific platform.
The research method, relying mainly on scenario research, proved to be useful but laborious. This was also noted by Ermi and Mäyrä in their earlier study with scenarios and user-centred game design (Ermi and Mäyrä 2005). The supportive research had definetly its place, as we were able triangulate, i.e. approach the design framework from multiple viewpoints.

Based on our experiences with current educational language games, we strongly suggest that designers take advantage from the current game design and evaluation guidelines presented in the related work section. If there are basic usability and playability problems in any type of games, the user will certainly frustrate and quite playing.
It became apparent that neither the teachers nor the pupils are satisfied with the current educational games. Closing the content quality gap between commercial leisure games and educational games might be impossible. Therefore we suggest that educational game design should focus on transforming traditional study book exercises into playful, game-like applications. Both short- and long-term extrinsic motivational factors are powerful, and the gap between playing an educational mini-game or doing a repetitive tasks in *World of Warcraft* to gain levels might not be as wide as expected. However, as Fortugno and Zimmerman state, both educators and developers must co-operate to increase the quality of educational games (Fortugno and Zimmerman 2005).

Table 2. The design framework for educational language games.

#	Title	Description	Example
#1	Context adaptation	The games should support different use contexts: class and home.	The game should support verbal social interaction in class environment, while at home the emphasis is on solitary tasks such as listening, reading and writing.
#2	Multi-layer game system	The game should be seen as a large system containing different modes for play.	The games are tied into the study book curriculum. Campaign games are featured in each chapter of the study book through the whole semester. Mini-games represent tasks that can be played out quickly at any time. There is also a persistent background virtual world, which the player can manage.
#3	Mechanics and learning	The game mechanics and learning should be intervowen.	The games require the skills that are needed when using a foreign language. Identifying contexts, social interaction, verbal communication, reading and writing skills. The play and learning are not seperated.
#4	Automatic differentiation	Automatic differentiation (adaptive difficulty) should be automatic and capable to adjust the difficulty to both directions.	Advanced pupils who perform well will receive harder tasks, where as slower learners receive easier tasks. Goal is to find the optimum difficulty level for each pupil, so that the challenge is in balance and learning is efficient.
#5	Multi-layer rewards	There should be instant, short-term and long-term rewards. The purpose of these rewards is extrinsic motivation.	The players receive different kind of points from the games. For example the mini-games produce instant score which can be use to customize the player's character. Campaign games produce points which can be used to build the virtual world. The last task of a campaing game, i.e. the last exercise of a chapter, features a more difficult task and succeeding will result to a special reward.
#6	Customizable character	There should be a virtual character, which is customizable.	The character should be always present and visible when playing any of the games. The character can be customized by using score received from the mini-games. The character is able to assist the player and perform simple interaction.
#7	Tried and true concepts	The games should be based on exercises, that have been found useful and motivating.	The mini-games feature exercises such as crossword puzzles and other vocabulary tasks. The campaign games feature more demanding tasks like listening comprehensions, grammar puzzles, group work in a class room, etc.

#8	Group work	Group work should be supported, when applicable.	In a class environment, the games should support working in pairs or in larger groups. Games could produce partial information for each member, who would then have to work together to finish a task.
#9	Performance measurement	Teacher should be able to monitor, evaluate, compare and report the pupil's performance.	The performance in games is logged into a server, which makes it possible to analyse the data and pinpoint problem areas that need attention.
#10	Short play sessions	15 minute play session for each game to keep the interest.	Play sessions for each game should not be too long, as the pupils will lose their interest quickly if the task is considered boring.

The virtual worlds and world building games surely have their place, especially when the short time drill and practice games are attached to them through reward mechanics for example. Although some authors, at least implicitly, consider drill and practice games old fashioned (Squire 2003), we believe that there should be both smaller and larger games, and also explorable open-ended worlds where the learning is not as explicit as with the other type of games. This kind of multi-layer game system, which would be connected to the study book curriculum, could be very captivating and motivating for the pupils to play both in school and at home.

The teachers considered that issues such as moral and ethics, which are difficult to teach according to the teachers, could be visualized in practice in virtual worlds, where the player sees the consequences of her actions. The downside of these game worlds is that they require a lot of time for pupils to comprehend. For example, it took 6-7 hours of game play for middle school students to understand the basic concepts in *Civilization III* strategy game (Squire 2005). If the pupils are motivated to play the educational games also at home, the burden to play at school would definetly decrease.

Combining educational games and hybrid media seems to be an interesting research field with lots of opportunities. Hybrid media is related to *ubiquitous computing* (e.g. Sakamura and Koshizuka 2005) which has gained a lot of attention from the research field. It has also shown the great diversity of different technological possibilities, which are also apparent with hybrid media. This fields requires more research on how these technological possibilities could be taken into advantage in pragmatic educational work.

FUTURE WORK

The *LehMa* project is part of a larger research effort in the area of learning combined with hybrid media. The field in which we believe there will be most obvious benefits of utilizing both printed and digital media side-by-side is language studies, especially in elementary schools. The role of book is strong, even indisputable in the time frame of a couple of decades. However, more convenient ways of merging the essential digital content (e.g. listening tasks) into everyday studies are needed.

Today schoolbook publishers provide CD's, internet sites and complementary digital material for teachers as supplementary material, but in practice, they are under-utilized due to the extra effort and time required for their use, both in school and at home. That said, if new hybrid products will not be brought to market in the near future, the most convenient solution might be a mini laptop or e-book, even though some unique advantages of paper would then be lost. In order to gain a more extensive understanding of the feasible alternatives in the development of learning materials, another project has been running side-by-side with the one presented here. In that project experts representing teaching, e-book technology, publishing and e-learning in Finland, UK and the Netherlands have been interviewed.

In the next phase of this project, hybrid school book prototypes have been prepared for user tests with elementary school children, their parents and their teachers. The results from this study have been used to create exercise examples for medium fidelity hybrid media platform and mobile phones. The prototypes have been prepared in the beginning of 2009, and the user tests will be carried out during summer 2009. Based on these user tests a final version of the hybrid study book will be designed and taken into actual class room use and evaluation in the autumn 2009.

ACKNOWLEDGMENTS

We would like to thank Mika Isomaa, Robert Roozeman, Antti Aikala and Timo Kuula for their ideas and efforts in this project. Special thanks for all the authors, teachers and pupils who were willing to be interviewed as you gave us invaluable feedback.

REFERENCES

Aguilera, M. and Méndiz, A. 2003. “Video Games and Education.” *ACM Computers in Entertainment* 1, No. 1 (Oct).

Annetta, L.A. 2008. “Video Games in Education: Why They Should Be Used and How They Are Being Used.” *Theory Into Practice, 47*. College of Education and Human Ecology, The Ohio State University. 229-239.

Bogost, I. 2007. Persuasive Games: The Expressive Power of Videogames. The MIT Press.

Desurvire, H., Caplan, M., and Toth, J.A. 2004. “Using Heuristics to Evaluate the Playability of Games.” CHI 2004, Vienna, Austria.

Ermi, L. and Mäyrä, F. 2005. “Player-Centred Game Design: Experiences in Using Scenario Study to Inform Mobile Game Design.” Game Studies 3, No.1 (Oct).

Federoff, M. 2002. “Heuristics and Usability Guidelines for the Creation and Evaluation of FUN in Video Games”. Master's Thesis. University Graduate School of Indiana University.

Fortugno, N. and Zimmerman, E. 2005. “Soapbox: Learning to Play to Learn – Lessons in Educational Game Design”. Gamasutra featured article, Apr 5, 2005. Available online: http://www.gamasutra.com/features/20050405/zimmerman_01.shtml

Gee, J. P. 2003. *What Video Games Have to Teach Us About Learning and Literacy*. Palgrave Macmillan, New York, NY, USA.

Gee, J. P. 2004. *Situated Language and Learning: A Critique of Traditional Schooling*. Routledge, London, UK.

Gee, J. P. 2005. *Why Video Games Are Good For Your Soul: Pleasure and Learning*. Common Ground, Melbourne, Australia.

Gee, J. P. 2007. *Good Video Games and Good Learning: Collected Essays on Video Games, Learning and Literacy*. Peter Lang Publishing.

Henriksen, T. D. 2006. Dimensions in Educational Game-Design - perspectives on designing and implementing game-based learning processes in the educational setting. Presented in Nordic Playground event in Reykjavik Apr 18 2006. Available online: http://www.dpb.dpu.dk/dokumentarkiv/Publications/20070104094423/CurrentVersion/Thomas_Duus_Henriksen___Dimensions_in_Educational_.pdf

Höysniemi, J., Hämäläinen, P., and Turkki, L. 2004. Wizard of Oz prototyping of computer vision based action games for children. In *Proceedings of the 2004 Conference on interaction Design and Children: Building A Community*. 27-34.

Kiili, K. 2005. “On Educational Game Design: Building Blocks of Flow Experience.” Doctoral Thesis, Tampere University of Technology publication 571. Tampere, Finland.

Korhonen, H. and Koivisto, E.M.I. 2007. ”Playability Heuristics for Mobile Games.” In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services* (Helsinki, Finland). ACM, New York, NY, 9-16.

Kämäräinen, A. 2003. “Opetuspelin käytettävyyden arviointi.” Master’s Thesis, University of Jyväskylä. Jyväskylä, Finland.

Malone, T. W. 1980. What makes things fun to learn? heuristics for designing instructional computer games. In *Proceedings of the 3rd ACM SIGSMALL Symposium and the First SIGPC Symposium on Small Systems* (Palo Alto, CA, Sep 18-19). SIGSMALL '80. ACM, New York, NY, 162-169.

Oittinen P. 2006. “Kehitysaskelia kohti hybridimediahyppyjä“. In *Hybridihenkselit metsäklusterin innovoinnon edellytyksiä*, T. Meristö, E. Torniainen and P. Kivistö, eds. 73-82. Faculty of Technology - Corporate Foresight Group CoFi / Åbo Akademi, Turku, Finland, 73-82.

Prensky, M. 2001. *Digital Game-Based Learning*. McGraw-Hill, New York, NY.

Sakamura, K. and Koshizuka, N. 2005. “Ubiquitous Computing Technologies for Ubiquitous Learning.” In *Proceedings of the third IEEE International Workshop on Wireless and Mobile Technologies in Education 2005* (Tokushima, Japan, Nov 28-30). IEEE, Piscataway, N.J., 11-18.

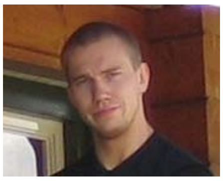
Seisto, A., Federley, M., Aarnisalo, S. and Oittinen, P. “Hybrid Media Application for Language studies in Elementary School.” *IADIS International Conference Mobile Learning 2009* (Barcelona, Spain, Feb 26-28).

Squire, K. 2003. “Video Games in Education.” *International Journal of Intelligent Simulations and Gaming* 2, No. 1.

Squire, K. D. 2005. “Changing the Game: What Happens When Video Games Enter the Classroom?” *Innovate* 1, No. 6.

Thomas, S. 2007. “Pervasive scale: A model of pervasive, ubiquitous, and ambient learning.” The Pervasive 2007 Workshops. Available online: http://www.massey.ac.nz/~hryu/PL2007/Siobhan.pdf

BIOGRAPHY



Janne Paavilainen is a project manager and game researcher at the Games Research Lab, University of Tampere, Finland. Janne has been in various research projects focusing on mobile games, game design and evaluation, educational games and social games. Currently Janne is focusing in *social playbility* in the SoPlay project (http://soplayproject.wordpress.com).

VISIBILITY CULLING USING OCCUPANCY PROPORTION AND ADPATIVE LOD ALGORITHM

Furen ZHENG*
Department of Animation School
Communication University of China
No.1 Dingfuzhuang East Street, Chaoyang District
Beijing, China
E-mail: zhengfr@163.com

KEYWORDS
Occlusion Culling, Occupancy Proportion, Adaptive LOD.

ABSTRACT

This algorithm uses the prioritized-layered projection to occlude those obscured objects and uses an approximate visibility technique to determine which objects will be visible in the coming future and prefetch those objects from disk before they are rendered. View-dependent rendering technique offers an efficient approach based on the new concept of occupancy proportion to visibility culling and provides the ability to change level of detail over the surface seamlessly and smoothly in real-time according to cell solidity value.

1 PREVIOUS WORK

In recent years, the subject of visibility culling of large environments has received substantial contribution. However the major amount of research into the area such as view-frustum culling, occlusion culling, back-face culling, contribution culling and portal culling. Our approach focus on occlusion culling which has a shortcoming that it is latency, our algorithm combines the advantages of PLP [1] algorithm, occlusion culling and contribution culling ,and exploits an efficient algorithm [2] for object visibility determination.
The paper is organized as follows: In section 1, the relevant work is discussed. In section 2, a novel visibility culling algorithm is proposed. The implementation details of our algorithm and corresponding experimental evidence is explained in section 3. In section 4, some future works and Occlusion culling is a technique whose goal is to determine which geometry is hidden from the viewer by other geometry, and is especially effective for scenes with high depth complexity. Starting with the pioneering work of Airey et al. [3] and Teller and Sequin [4]on precomputing lists of potentially visible geometry from cells, visibility culling grew into an important area of research in computer graphics. Now some classification of occlusion algorithms: image-precision algorithm [5,6], object-precision algorithm [7,8], from-region algorithm [9,10] from point algorithm [11]. In Addition, Samuli et al.[12] exploited directional coherence of visibility between adjacent view cells to get output-sensitive visibility preprocessing, Oliver presented a new algorithm for efficient occlusion culling using hardware occlusion queries [13],Deb presented an efficient algorithm for visibility culling that supports static and dynamic scenes with equal ease with significant performance improvements over existing schemes [14], Mansa proposed a small addition to the GPU occlusion queries to perform faster renderings [15],and used occlusion queries and succeeded in improving the frame rate of the stereoscopic visualization with an effective scheduling of the queries [16].We refer the interested reader to the recent surveys in [17,18]. Our approach provides conservative and approximate selection for different application.

2 OUR APPROACH

Reducing the number of triangles sent to graphics hardware accelerates the rendering of polygonal datasets. To address this problem, solutions such

*、Animation School， Communication University of China, Tel 8610-65779279,E-mail:zhengfr@163.com.

as occlusion culling, geometric simplification, occlusion fusion etc. are developed.
In some situation, some approaches can obtain higher rendering efficiency than conservative object bounding box occlusion query method. Our approach integrates this algorithm [19] well.
In some situation, maybe only 1 percent of the geometry is visible but it is rendered in high LOD in that it is very close to current viewer. Our approach integrates these two fundamental approaches: occlusion culling and LOD rendering. According to the user's various demands, our system can deal with different situation well. Our Rendering System focuses on four parts of core algorithm marked with I, II, III, IV.

2.1 Part I: Occulusion CullingSecondary.
In this part, PLP algorithm is used to compute an approximate visible data set, and clip [20] is used to compute a conservative visible data set.
To understand our algorithm, PLP algorithm introduced by Klosowski and Silva will be reviewed below. It's an approximate occlusion culling technique. Firstly, instead of performing conservative visibility determinations, PLP is an approximate occlusion culling algorithm that estimates the visible geometry for a given viewpoint, and renders those primitives who it determines to be most likely visible. Second, instead of traversing the entire hierarchy, PLP works on a user-defined budget, it stops the traversal after lots of primitives have been determined to be added to the visible set.
Fig.1 shows our rendering system architecture. For different camera parameters, the system recomposes the primitives have been determined to be added to the visible set. Thirdly, a cell containing much geometry is likely to occlude the cells behind it. PLP maintains a priority queue called the *front*, which determines which cell is most likely to be visible. When PLP visits a node, it adds it to the *visible set*, removes it from the *front*, then adds the unvisited neighbors of the node to the *front*. Fig.2 describes PLP algorithm's RenderingLoop briefly.

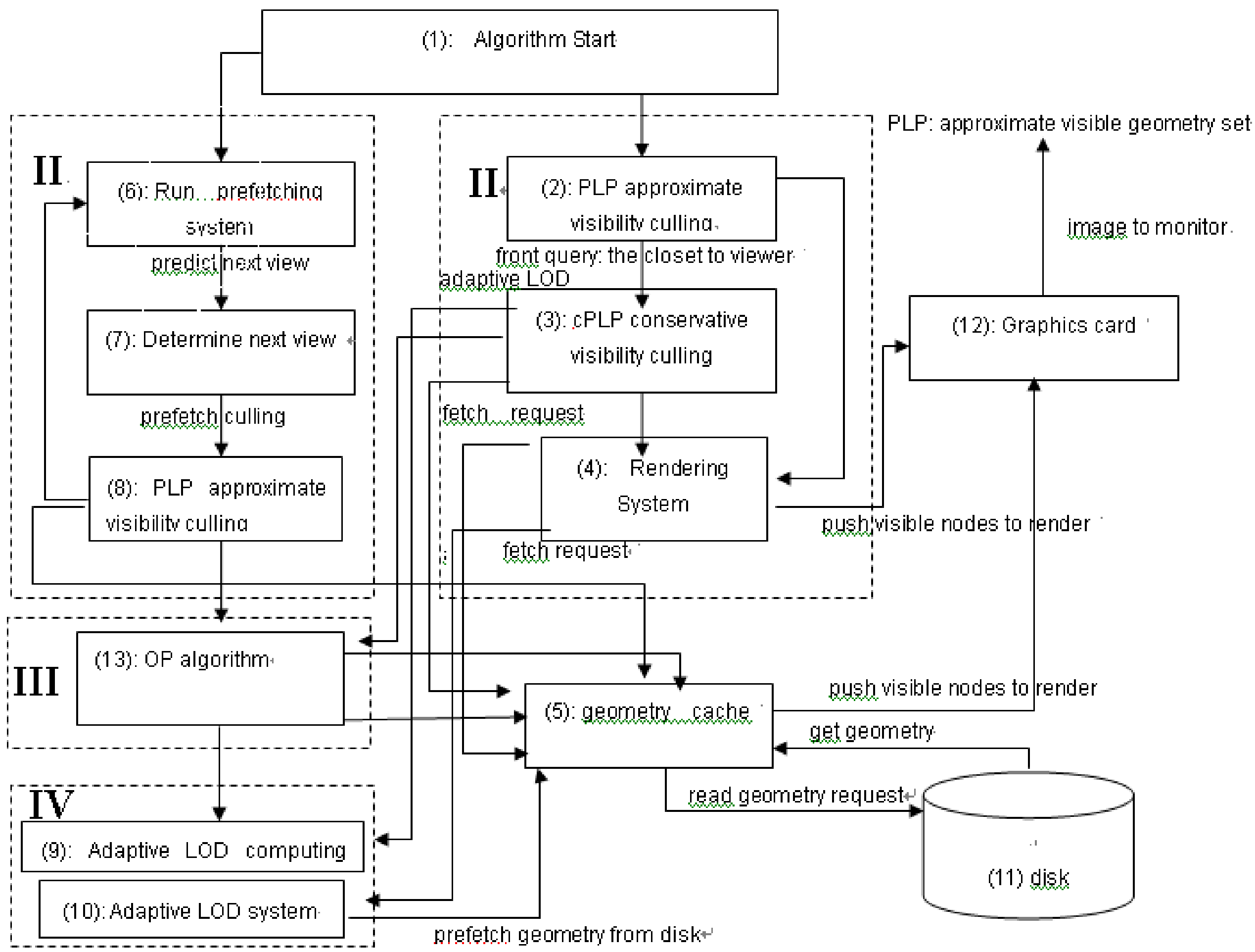


Figure 1: Algorithm system architecture

Similarly, conservative Prioritized-Layered Projection (cPLP) is a conservative extension of PLP, which guarantees 100% accurate images. Compared with PLP algorithm, cPLP has he following characteristics:

① 100% image quality

② Slower than PLP algorithm because of additional disk operations

③ Using item-buffer technique to be portable to any platform that supports OpenGL or occlusion query.

Fig. 2 shows the pseudo-code for this.

```
ConservativeFinish (RenderingContext)
    select  occlusion query mode
    if mode is OPENGL
         ConservativeFinish_GL(RenderingContext);
        if mode is HP Occlusion Query
   ConservativeFinish_HP(RenderingContext);
        if mode is nVidia Occlusion Query
          ConservativeFinish_NV(RenderingContext);
        else return
```

Figure 2: Pseudo-code for Item_buffer Technique

2.2 Part II: Prefetching Algorithm.

In the past decades, many researchers have studied prefetching approach. From-region visibility algorithm is studied in [20,21] etc. From-point algorithm is studied in [17] and so on. Our prefetching method, from-point visibility algorithm is based on PLP algorithm. Firstly, we use a separate thread to deal with disk operations and communicate with PLP or cPLP algorithm thread etc. Secondly, the PLP is used to quickly compute the approximate visible sets, after our prefetching method predicts next viewpoint, prefetching system calls PLP algorithm to get approximate primitives in the near future, then pushes these approximate visible sets to priority queue *front* .

In Fig.3, prefetching method is shown.

```
Prefetcher(CurrentView)
        prefetch closest leaf's neighbors
        if  GetClosestLeaf() is not NULL
   prefetch GetClosestLeaf()->GetNeighbors()
   computing next view point
   define primitive set V;
   define primitive set F;
   RunPrefetchAhead(V, F, CurrentView);
   prefetch visible set: prefetch(V);
   prefetch front: prefetch(F);
```

Figure 3: Pseudo-code for Prefetcher

2.3 Part III: Occupancy Proportion Algorithm.

Our algorithm uses occupancy proportion (OP) technology which represents the tightness of the object bounding box. We issue the bounding box of the object instead of the object. In order to increase rendering occlusion query is usually issued for object bounding box, and the test is performed by retrieving the object bounding query result for the object. Actually issuing object bounding box occlusion query instead of issuing object is a conservative method. If the object is visible the box occlusion query result above 0, if the object is actual not visible the box occlusion query result can also above 0 possibly or is 0. Tightness is one of criterions to evaluate object bounding box efficiency.

Bounding box tightness can be represented in Hausdorff distance:

$$t = \max_{b \in B} \min_{g \in G} dist(b, g)$$

where t is Hausdorff distance, B represents object bounding box and G represents the responding object. If Hausdorff distance is more 1, the bounding box is more efficient. In actual scene, some object bounding boxes are not tight; it will render objects which are actually not visible. Hausdorff distance which presents the tightness of bounding box is mainly on the basis of geometry space, and it is also troublesome to be calculated.

Like Hausdorff distance we put forward another method to present the tightness, which is occupancy proportion (OP). OP is an approximate method and based on image space. The concept of object OP is based on directional OP. Directional OP is the ratio between object visible pixel number and its bounding box visible pixel number in 2D screen image space in a given direction for orthogonal projection. In our algorithm the bounding box adopted is Axis Align Bounding Box (AABB), and to an AABB we can query and get the visible pixel numbers (B_x B_y B_z) in 3 view directions which are parallel to 3 main x axis, y axis, z axis. Similarly we can query and get the object visible pixel numbers (O_x O_y O_z) in the same 3 view directions. Three directional OP (P_x P_y P_z) can be calculated by:

$$Px = \frac{Ox}{Bx}, Py = \frac{Oy}{By}, Pz = \frac{Oz}{Bz}$$

Object OP is the maximum of the three directional OP (P_x P_y P_z). Lastly, we select the maximum value among the three directional OP as object OP value:

$$P = \max(Px, Py, Pz)$$

2.4 Part IV: Adaptive LOD.

An approach to avoid assigning a high resolution to close primitive is invisible from current viewpoint.

In order to better introduce Adaptive LOD, PLP's core part will be reviewed as follows.

In Fig.4, skeleton of RenderingLoop (PLP core traversal algorithm) is described.

```
Rendering Loop ()
  while (empty(F)!=true)
    c = min(F)
    project(c)
    if ((reached_budget() == true)
        break;
      for each n; n = cell_adjacent_to(c)
        if ((projected(n) == true)
          continue;
        p = update_solidity(n,c)
      enqueue(n,p)
```

Figure 4: Pseudo-code for RunPrefetchAhead

① Function $\min(front)$ returns the minimum element in the priority queue $front$.

② Function $project(c)$ renders all the elements assigned to cell c.

③ Function $reached\_budget()$ returns true if PLP has rendered k primitives.

④ Function $cell\_adjacent\_to(c)$) lists the cells adjacent to cell c.

⑤ Function $project(n)$ returns true if cell n has already been projected.

⑥ Function $update\_solidity(n, c)$) computes the updated solidity of cell n . c is one of its neighbors just been projected. If n was already in the queue, this faction will remove it from the queue, and reinsert it with the updated solidity value.

The notion of a cell's solidity (between 0.0 and 1.0) is at the heart of RenderingLoop algorithm shown in Fig.3, and is related to how difficult it is for the viewer to see a particular cell, cells are removed from the queue $front$ in solidity order, PLP uses a set of conditions to roughly estimate the visibility likelihood of the cell and makes sure that cells more likely to be visible get projected before cells that are less likely to be visible. The heuristic to define the solidity value of cells is shown in Fig. 5.

In Fig.5, ρA is the accumulated solidity value for cell A. ρB is the solidity value to be computed for cell B.

Now what Part III will contribute to rendering system is that we determine the level of detail at a node by computing its view parameters contribution to estimate its occlusion probability. In a word, we integrate occlusion culling approach and LOD rendering approach.

```
update_solidity (B,A)
    ρB = |A| / ρmax + (v • nB) * ρA
    if((star_shaped(v, B) == false)
        ρB = apply_penalty_factor(ρB)
    return ρB
```

Figure 5: Update_solidity in PLP

The probability value can reduce the LOD at a node from the LOD computed by the view parameters to the lowest possible one. The Equation (1) is below:

$$LOD0 = \rho B * LOD1 + (1 - \rho B) * LOD2 \quad (1)$$

In Equation 1, $LOD1$ is the lowest possible LOD and $LOD2$ is current LOD computed by view parameters. ρB is solidity (PLP occlusion probability) at the cell. We use Equation 1 to get the final LOD, that is $LOD0$.

3 IMPLEMENTATION DETAILS

We have implemented our rendering system described above on two machines, machine 1 with Red Hat 9.0 operating system, with 2.0GHZ, GeForce graphics board, 1024 MB of main memory, machine 2 with Red Hat 7.0 operating system, with 1.7GHZ, GeForce 3 graphics board,

model	triangles	points	data
Bunny	69451	35947	2.2MBit
Lucy	14,027,872	28,055,742	508 MBit

Table 1: Bunny models and Lucy models.

256 MB of main memory. Our system uses C++, GLUT, and OpenGL.

Our algorithm is tested on two large environments, a bunny model and lucy model. Tab.1 tabulates the two models' triangles count, number of points, and size of scene graph data.

Firstly, we run PLP and cPLP algorithm on these two large environments, all the testing above needs view path at first. We specify only one path for all the tests.

In Fig.6, there are four situations as follows:

(a).Our rendering system runs on machine 1, the rendering frame number per second (fps) with prefetching or not prefetching is similar in that the bunny model is small compared with main memory size. If we run PLP with adaptive LOD algorithm, rendering system is optimized sharply. In order to test adaptive LOD algorithm, we make three bunny models with different LOD firstly.

(b).In this part, the data size of Lucy model is 508MB, over main memory, so without prefetching technique, the red curve with slot shows that the image coherence is not satisfied. What's more，our system can render those models requiring the memory with prefetching technique by our system.

(c).Compared with (a), rendering speed is faster because of machine2' configure is better than machine1.

(d).Compared with (b), the slot disappears because machine 2 can loads models without burden because of memory limit.

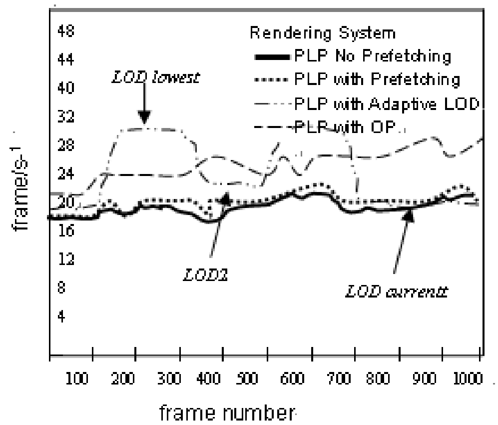
4 CONCLUSIONS AND FUTURE WORK

In this paper, our novel approach combines occlusion culling algorithm with adaptive LOD algorithm successful. Our system is tested on different machines with different input models.

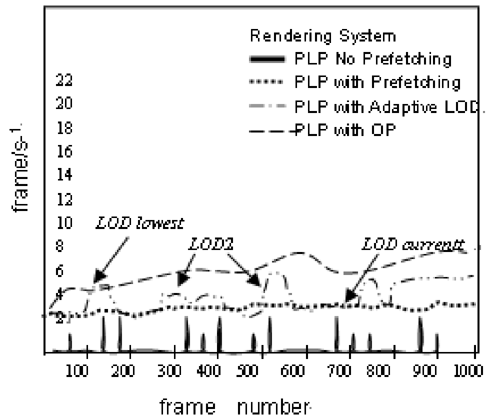
There are several possible areas for future work. One is to combine our algorithm with other useful algorithm such as occlusion fusion technique, distributed computing for rendering large models on machines [22] etc., to extend the system to support dynamics scenes [12], now out rendering system is only used to compute static LOD.

ACKNOWLEDGMENTS

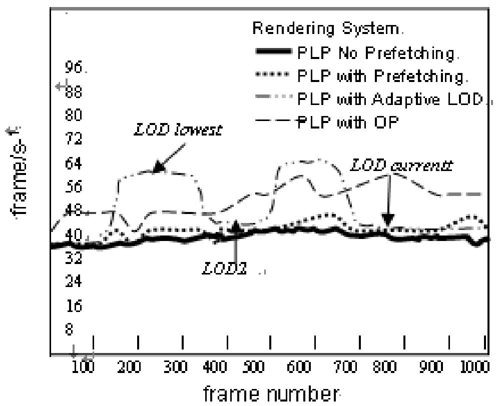
Our work was supported by 3rd "211Project" of Communication University of China



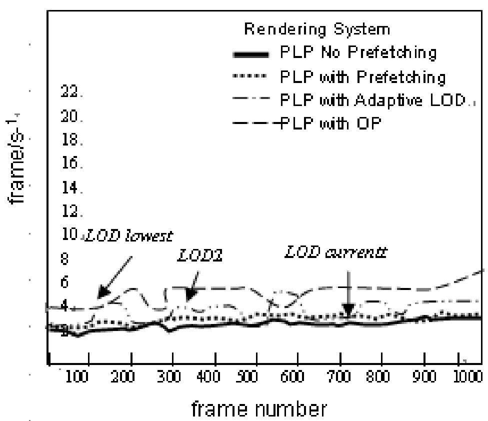
(a): Bunny Model Rendering on Machine 1



(b): Lucy Model Rendering on Machine 1



(c): Bunny Model Rendering on Machine 2



(d): Lucy Model Rendering on Machine 2

Figure 6: Bunny model and Lucy model rendering on two different machines.

(no.21103040204).

REFERENCES

[1] James T. Klosowski and Cláudio T. Silva. 2000. "The prioritized-layered projection algorithm for visible set estimation". In *IEEE Transactions on Visualization and Computer Graphics.* IEEE, Piscataway, N.J., USA, 108–123.

[2] S. Nirenstein et al. 2002. "Exact From-Region Visibility Culling". In *Proceedings of the 13th Eurographics workshop on rendering* (Pisa, Italy). Eurographics Association Aire-la-Ville, Switzerland, 191-202.

[3] John M. Airey et al. 1990. "Towards image realism with interactive update rates in complex virtual building environments". In *Proceedings of the 1990 symposium on Interactive 3D graphics* (Snowbird, Utah). ACM, New York, NY, USA, 41-50.

[4] Teller S. and Sequin C. H. 1991. "Visibility preprocessing for interactive walkthroughs". In *Proceedings of the 18th annual conference on Computer graphics and interactive techniques.* ACM, New York, NY, USA, 61-70.

[5] Greene N. et al. 2001. "A quality knob for non-conservative culling with hierarchical z-buffering". Technical report. In ACM SIGGRAPH Course Notes.

[6] Zhang H. et al. 1997. "Visibility culling using hierarchical occlusion maps". Technical report TR-97-004. University of North Carolina at Chapel Hill Chapel Hill, NC, USA.

[7] Satyan Coorg and Seth Teller. 1997. "Real-time occlusion culling for models with large occluders". In *Proceedings of the 1997 symposium on Interactive 3D graphics*. ACM, New York, NY, USA, 83-ff.

[8] Zheng Furen and Zhan Shouyi. 2003. "Applied Clipping Based on Multi-bounding Types in SVRT". In *Fourth International Conference on Virtual Reality and Its Applications in Industry* (Tianjin, China, Dec.23). SPIE, Bellingham WA, ETATS-UNIS, 87-91.

[9] Durand F. et al. 2000. "Conservative visibility preprocessing using extended projections". In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques.* ACM, New York, NY, USA, 239-248.

[10] Schaufler G. et al. 2000. "Conservative volumetric visibility with occluder fusion". ". In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques.* ACM, New York, NY, USA, 229-238.

[11] Sung-Eui Yoon and Brian Salomon. 2003. "Interactive View-Dependent Rendering with Conservative Occlusion Culling in complex Environments". In *Proceedings of the 14th IEEE Visualization 2003 (VIS'03).* IEEE, Washington, DC, USA, 22.

[12] S. Laine. 2005. "General Algorithm for Output-Sensitive Visibility Preprocessing". In *Proceedings of the 2005 symposium on Interactive 3D graphics and games* (Washington, District of Columbia,). ACM, New York, NY, USA, 31-40.

[13] Oliver Mattausch et al. 2008. "CHC++: Coherent Hierarchical Occlusion Culling Revisited". In *Computer Graphics Forum (Proceedings Eurographics 2008).* Blackwell Publishing ,221-230.

[14] Deb, S. and Gupta. 2008. "Visibility Cuts: A System for Rendering Dynamic Virtual Environments". *In ICVGIP Sixth Indian Conference on Volume* (Indian, Dec.16-19). Graphics & Image Processing, 47–54.

[15] Mansa and Amundarain. 2006. "Towards Adaptive Occlusion Culling Using Camera Coherence". In *Proceedings of the conference on Information Visualization*. IEEE, Washington, DC, USA, 591-596.

[16] I.Mansa et al. 2007. "Occlusion culling for stereoscopic display of complex mechanical sets". In *Proceedings of the 11th International Conference Information Visualization*. IEEE, Washington, DC, USA, 786–791.

[17] Daniel Cohen-Or et al. 2001. "Visibility, problems, techniques and applications". In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques.* ACM, New York, USA.

[18] Daniel Cohen-Or and Yiorgos Chrysanthou. 2000. "A survey of visibility for walkthrough applications". *IEEE Transactions on Visualization and Computer Graphics*, Volume 9, Issue 3, July-Sept. Page(s):412-431.

[19] Bo Li et al. 2008. "Efficient Occlusion Culling with Occupancy Proportion". In *International Conference on Computer Science and Software Engineering*. IEEE, 1058–1061.

[20] Klosowski T. and Silva C. T. 2001. "Efficient Conservative Visibility culling Using The Prioritized-Layered Projection Algorithm". *IEEE Transactions on Visualization and Computer Graphics*, Volume 7, no.4, Page(s):365-379.

[21] Thomas A. Funkhouser. 1996. "Database Management for Interactive Display of Large Architectural Models". In *Proceedings of the conference on Graphics interface '96*. Canadian Information Processing Society Toronto, Ont., Canada, Canada, 1-8.

[22] Greg Humphreys et al. 2001. "WireGL: A Scalable Graphics System for Clusters". In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques.* ACM, New York, NY, USA, 129-140.

BIOGRAPHY

Dr. Furen Zheng is a Research Staff Member in Virtual Reality and received his M.Sc. (2003) and Ph.D. (2005) degrees from Beijing Institute of Technology 's Department of Computer Science, and joined Animation School of Communication University of China in 2005.

His research interests include virtual reality, large-scale machine learning, game design and data mining. Dr. Zheng has authored or co-authored 2 book chapters and more than 10 international conference and journal papers.

USING CUDA
IN XNA GAMES DEVELOPMENT

Dar-jen Chang and Ming Ouyang
Department of Computer Engineering and Computer Science
University of Louisville
Louisville, Kentucky 40292
E-mail: djchan01@louisville.edu

KEYWORDS
Graphics processing unit (GPU), GPU programming, CUDA, game engines, XNA.

ABSTRACT

Graphics processing units (GPUs) are powerful computational devices tailored toward the needs of the 3D gaming industry for high-performance, real-time graphics engines. Since 2006, Nvidia has released several generations of GPUs designed for graphics as well as general-purpose computing. These GPUs have more number of processing cores than CPUs. For example, G80 series GPUs provide 128 general-purpose stream processors (SPs) and support 12,288 active threads. And more recent GTX200 series GPUs have 240 SPs and support double-precision floats. These capabilities allow for the design of innovative, non-graphics application for GPUs such as database searching, protein folding, and physics simulation. In 2007, Nvidia released a general-purpose GPU programming language called CUDA, which is an extension to the C programming language. On the game development side, XNA provides a powerful and friendly C# game programming environment for Windows, Xbox 360, and Zune games development. Many games can take advantages of GPU computing capability not only for graphics rendering but also for computationally intensive tasks such as game physics and game AI. In this paper we intend to bring the awareness of potential benefits of using GPU in games for non-graphics tasks. Specifically we describe how CUDA can be used in XNA game development for general-purpose, computationally intensive tasks.

INTRODUCTION

Graphics processing units (GPUs) on commodity video cards have evolved into powerful computational devices tailored toward the needs of the 3D gaming industry for high performance, real-time graphics engines. In the past, software development on GPUs has been geared exclusively toward graphics through the use of shader languages such as OpenGL shading language (GLSL) and Direct3D high-level shader language (HLSL). Since 2006, Nvidia has released several generations of GPUs designed for graphics as well as general-purpose use. For example, G80 series GPUs have up to 128 stream processors (SPs) and support 12,288 active hardware threads. And more recent GTX200 family GPUs have up to 240 SPs and support double-precision floats. This new GPU architecture facilitates efficient general-purpose computing on GPUs (GPGPUs). In 2007, Nvidia released an extended C language for GPU programming called CUDA [1], short for Compute Unified Device Architecture. Using CUDA, innovative data-parallel algorithms can be implemented in general computing terms to solve many important and non-graphics applications such as database searching and sorting, medical imaging, protein folding, and physics simulation.

On the game development side Microsoft XNA Game Studio [2] enables developers to create video games for Windows, Xbox 360, and Zune by using optimized cross-platform gaming libraries based on the .NET framework and the C# programming language is commonly used to write XNA games. Shaders written in HLSL have been routinely used in XNA games for real-time graphics rendering on GPU. However, there are many non-graphics tasks such as game physics and game AI, which are too computationally intensive to be run on CPU in real time and difficult to be written in HLSL to run on powerful GPU. CUDA opens up a familiar language approach for game developers to apply more realistic game physics and sophisticate game AI. We will describe how CUDA can be used in XNA games development.

XNA GAME LOOPS AND CUDA

XNA consists of several cross-platform gaming libraries based on the .NET framework to support common game programming tasks ranging from game loops, graphics, sound and video playback, networking, to content pipeline. To use CUDA in XNA games development we have to understand how XNA game loops work. Simply put, a game loop is a loop of an *update* call following by a *draw* call repeating until the game exits. A game loop can use either fixed-time step or variable-time step, but in either case a loop has to be completed in less than 33 ms (i.e. for 30 frames per second) to ensure a smooth game play. GPU shaders written in HLSL are commonly used in the *draw* call to speed up graphics rendering while to achieve high graphics quality and special visual effects. The *update* call normally is used to update the game states such as the positions, velocities, and decisions of game entities due to the game physics or game AI constraints. This *update* is commonly done in CPU or custom chips for physics or AI. In many cases CPU is too slow and custom chips are too difficult if not impossible to program for any desired game update tasks. To take advantages of powerful GPUs programming in HLSL probably is as difficult as using custom chips for non-graphics applications.

The video cards based on Nvidia G80 and GTX200 family GPU can be used for dual purposes of display and CUDA computing, whereas the CUDA-enabled cards with no display circuitry such as Nvidia Tesla are only used for CUDA computing. This difference in using GPU has to be taken into account when applying CUDA to XNA games since dedicated CUDA cards cannot interoperate with the graphics engines (OpenGL or Direct3D, for example) directly.

CUDA DRIVER API

CUDA is a general-purpose parallel computing architecture – with a new parallel programming model and instruction set architecture – that leverages the parallel computing capability in CUDA-enabled devices to solve many complex general-purpose computational problems in a more efficient way than on a CPU. CUDA comes with a software environment that allows developers to use extended C as a high-level programming language. The CUDA software environment includes a compiler for translating CUDA GPU code into GPU device machine code, which can be loaded into GPU memory for execution. The lowest level of CUDA interface to GPU devices is the CUDA driver API (application programming interface), whose functionality is exposed through a C library. The driver API provides a comprehensive interface to CUDA devices but it is rather tedious to use for CUDA applications. The C for CUDA in the CUDA distribution is an easier-to-use C library built on top of the driver API and most of the demo projects in the CUDA distribution use this C library.

The driver API includes functions to manage devices, contexts, streams, events, modules, kernel functions and GPU memory. A typical programming flow of the driver API programming for CUDA applications is as follows:

1) Call `cuInit()` before calling other deriver functions.
2) Get a CUDA device.
3) Create a context (context is kind of GPU process) for the device.
4) Allocate GPU memory and copy host data to GPU memory.
5) Load modules and get functions from loaded modules.
6) Set function parameters.
7) Set thread block dimension.
8) Launch (execute) kernel functions with a specified thread block grid dimension.
9) Copy the result from GPU memory to CPU memory.

Additionally, programmers can call driver API to dynamically allocate shared memory to thread blocks and use texture hardware available in CUDA devices. Because G80 GPU can yield about 388.8 GFLOPS (and about 729 GFLOPS for GTX200) of peak theoretical computing performance and global memory access takes significantly more machine cycles (between 400 to 600 cycles) than that of operations, it is important to design computing algorithms to hide the global memory latency to remove the memory wall that prevents the devices to achieve high throughput.

USING CUDA IN XNA GAMES

Since the CUDA driver API exposes its functionality through a C library, it is very easy to write a desired language interface to the driver API. To use CUDA in XNA games programming, we need a C# interface to the driver API. GASS CUDA.NET [3] is an effort to provide access to CUDA functionality through .NET applications. The current version (2.2) of GASS CUDA.NET supports CUDA 2.2 and can be used with C#, VB.NET, managed C++, etc. GASS CUDA.NET includes a CUDADriver class which contains static methods that are one-to-one correspondence to the driver API functions (with some adjustments of data type differences between C and C#). Moreover, GASS CUDA.NET also includes a CUDA class, which uses a better object-oriented approach and provides a simplified interface (similar to C for CUDA in the CUDA distribution), and classes to use BLAS (linear algebra) and FTT (fast Fourier transforms) libraries from the CUDA distribution.

We have tested the GASS CUDA.NET's CUDA and CUDADriver classes in some XNA 3.0 Windows game projects using Visual Studio 2008. (Note Xbox 360 and ZUNE do not support CUDA.) We have found that using the CUDA class is simpler in device memory allocation and data transfer between CPU and device memory than using the CUDADriver class. Other than that, the CUDADriver class is as easy to use as the CUDA class. As having said before, the CUDA driver API provides a comprehensive control of the devices, some of whose functions may not be available to higher-level APIs such as C for CUDA and the CUDA class.

To use CUDA for XNA games, we have to make sure that the design of computing algorithms takes into account of the strengths and weaknesses of CPU and GPU, respectively. Some tasks may be more suitable for CPU, while others may be more suitable for GPU. For GPU processing, data have to be transferred from CPU memory to GPU memory. This is the overhead not present in CPU processing. For systems with one video card used for both display and CUDA computing, there is less device memory to use for CUDA (the display uses the same 32-bit addressing space) and CUDA computing shares the computing cores with the display shaders. However, this dual-purpose video card allows interoperability between CUDA and graphics engines (i.e. OpenGL and Direct3D). In other words, CUDA can access vertex or pixel buffers used for graphics rendering. This interoperability may save trips to transfer data between CPU and GPU memory. For systems with separate video (display) card and CUDA-enabled card (such as Tesla), there is more device memory for CUDA computing. As far as we know these systems cannot transfer data between the video card and CUDA device memory directly. Data must be transferred to CPU first and then transferred to the destination memory.

In the XNA game loop, for non-graphics updates, we can insert CUDA call in the *update* method to do the job. In some non-real time applications, we need to achieve tens of speedup of using CUDA over using CPU to justify the CUDA development cost. However, in real-time games any speedup is significant since there is such a small time window for game update in each game loop. This is the main reason why CUDA is very useful for game applications. XNA games usually have a very spread-out object-oriented object structure (e.g. each entity object contains its own data), whereas CUDA computing tends to rely on linear memory structures for all object data. This disparity in data structures may require data conversions between game data structures and CUDA data structures before CUDA kernels are launched. Thus we must carefully weigh the added memory overhead and data conversions time with the gain of CUDA speedup.

POTENTIAL CUDA APPLICATIONS IN XNA GAMES

Any non-graphics tasks that are computationally intensive (e.g. nonlinear, polynomial time complexity algorithms) can be targeted for CUDA speedup. We think there are two main areas of game tasks which can be considered for CUDA speedup, namely, game physics and game AI.

Nvidia PhysX [4] is a game physics engine claimed to use CUDA. However, we do not have detailed information about how CUDA is used to accelerate PhysX and testing data of CUDA (say running on a Tesla card) speedup over CPU or PhysX PPU (Physics Processing Unit). But it is obvious that there are many game physics problems beyond those supported in PhysX. So it is worthwhile to devoting research and development effort in CUDA for game physics.

Game AI may also require intensive computing resources. For example, the flocking steering behaviors of game agents [5] essentially involve pair-wise interaction of these game agents. If there are N agents, the time complexity will be N^2. We have tested a simple flocking behavior involving around 2000 agents. The CPU implementation simply would not run in real-time, whereas the CUDA implementation can run in real-time for 64K agents on a Tesla C1060 card. Path finding is another important, computationally intensive game AI problem. The path finding problem can be formulated in terms of graph search. CUDA is very useful and straightforward for searching large-scale, densely-connected graph using an adjacent matrix data structure. For sparsely-connected graphs, an adjacent lists data structure is more memory-efficient than an adjacent matrix. In such a case, care has to be taken in CUDA implementation to make sure that the adjacent lists data structure would not require too many un-coalesced global memory accesses, which are rather expensive in current CUDA device architecture.

FUTURE WORKS

For future works, we intend to apply our CUDA knowledge we learned from testing the CUDA implementation of the flocking steering behaviors to other game AI problems. These game AI problems include path finding, tactics and planning, learning and adaptation, and so on [6].
It would be nice to have a CUDA-enabled game AI engine. We will also investigate applications of Multi-GPU to games and deal with the 4 GB GPU memory limitation problems (current Nvidia GPU only supports 32-bit addressing space).

REFERENCES

[1] Nvidia Corporation CUDA Zone http://www.nvidia.com/object/cuda_home.html
[2] XNA Creators Club http://creators.xna.com/
[3] GASS CUDA.NET http://www.gass-ltd.co.il/en/products/cuda.net/
[4] Nvidia Corporation PhysX http://www.nvidia.com/object/physx_new.html
[5] Mat Buckland, *Programming Game AI by Example*. Wordware Publishing, Inc., 2005.
[6] Steve Rabin, Ed, *AI Game Programming Wisdom 4*, Course Technology, 2008.

BIOGRAPHY

Dar-jen Chang is an associate professor in Computer Engineering and Computer Science at the University of Louisville. His research interests include computer graphics, game engines, scientific visualization, GPGPU and CUDA, and 3D modeling.

Ming Ouyang is an assistant professor in Computer Engineering and Computer Science at the University of Louisville. His research interests include bioinformatics, high-performance computing, GPGPU and CUDA.

GROUND COVER AND VEGETATION IN LEVEL EDITORS: LEARNING FROM ECOLOGICAL MODELLING

Eugene Ch'ng
School of Computing and Information Technology
The University of Wolverhampton, Wulfruna Street, WV1 1SB, Wolverhampton
email: e.chng@wlv.ac.uk

ABSTRACT

Ground cover and vegetation are important elements in large outdoor gaming environments. These include game scenes such as forest terrains, rural and urban environments. While the graphical representations of plants may look unbelievably real, a crucial factor related to their placements often lacks a natural association with those seen in the physical world. This article explores and pilots a study for simulating vegetation dispersal patterns and object positioning for game scene creation. The simple algorithms simulate growth patterns similar to nature and demonstrate the potentials of application in game level editors.

Keywords: Vegetation, ground cover, Scene Design

1. Introduction

Natural landscapes and large scale rural and urban environments have become a common scene in modern 3D computer game productions. While the look and feel of natural objects such as trees and rocks may create in players a sense of reality or presence (Minsky 1980) – the sense of being there, their functional cause are of much more importance to the survivability of the player. Trees and natural barriers could act as protection, camouflage, resources and even as weapons.

The impressive natural landscapes seen in successful First Person Shooters (FPS) such as Ubisoft's FarCry series and Electronic Art's Crysis are seminal in outdoor game scene creation. Such high quality environments warrant high requirements from the artists and designers that created the games. However, not all artists or designers are equal and it may be difficult to find designers with the botanical knowledge of 'what grows where'. There may be advanced developments in the behaviours of Non-Player Characters (NPC) and Artificial Life algorithms for bird flocks (Reynolds 1987; Lecky-Thompson 2008) inspired from nature, but gaps exist where such developments could be benefited by the inclusion of simple models inspired from the sciences such as ecological modelling (Gillman and Hails 1997). If algorithms can be adopted from this particular research area, creating vegetation covers for outdoor scenes would be made simpler. Artists need only 'block-in' a selection of vegetation onto the landscape without having to reference pictures of nature nor know the niche of plant types in differing global and local environments (Grime, Hodgson et al. 1988) such as tropics and seasonal plants, and their preferences for soil types, water logging, and altitudinal limits among the few. Scene designers need only create an environment conducive for plant growth using heightfield-based soil types and user interface sliders for setting climates and plant preferences. Current game level editors as we shall soon see, lacks these features. The integration of such features will certainly enrich the pre-production process of games.

This article describes the techniques used for piloting a study related to 'growing' vegetation inspired from a field of ecological modelling for game scene creation. Section two briefly explores the background related to this research. Section three describes the techniques and algorithms used for simulating simple vegetation growth. Section four demonstrates the simulation processes in different game scenarios and section five concludes the paper with plans for future research.

2. Background

WYSIWYG 3D FPS level editors first became useful when UnrealEd (UEd) accompanied Epic's Unreal game. The inclusion of a level editor greatly extends the lifespan of the game, allowing players to bridge the gap between inexpensive content generation tools such Autodesk's 3D Studio Max and the now outdated Gmax and a game MOD. Furthermore, the increasing requirements for realistic outdoor scenes initiated the integration of vegetation and object 'painting' tools within these editors (E.g., CryEngine Sandbox Editor I & II) so much that it has now become a mandatory feature in games engines (E.g., Torque 3D Game Engine, Unity3D Engine).

Vegetation and object placement tools differ among games engines. Some, like CryEngine's sandbox editor allows parameter tuning for density, random scale, swaying, plants and object models selection before they are 'painted' onto the landscape. Others like Torque 3D Game Engine's Foliage Replicator and Shape Replicator allows bitmap images of plants and 3D models to appear within a shape boundary (an ellipse) with settings for viewing distance, density, and dimensions. So far, although such features will extensively automate plant placements on a landscape, not one level editor incorporates ecological features for defining vegetation niche related to their adaptability on environments of variable conditions. Although proprietary terrain generators such as L3DT, Digital Element's WorldBuilder (McClure 2009) and 3D Nature's World Construction Set and Visual Nature Studio (Huber 2004) include certain features that allows automated placements of vegetation (terrain convexity/convavity and direction facing slope), more could be developed to extend the functionality to include the creation of different plant types and how each species could be associated with environmental factors that allow them to be distributed naturally across a terrain. L3DT possesses partial environmental attributes such as water flooding, water-table, salinity map, and attributes map that affects the texture of the terrain but do not generate plant to terrain positioning. The lack of exportable 3D coordinates of plants and objects in association with terrain size makes it difficult when large terrains are needed.

In order to create a natural looking terrain, scene designers frequently reference pictures of natural landscape when placing vegetations and terrain objects. This works for manageable sized terrain and outdoor scenes, but when large scale terrains and complex scenes are needed, their job becomes drastically tedious. Although, level editors have features that provide various parameters which designers can tune to set the scale, rotation, and position, tree types and density, but requires careful 'painting', or blocking in a community of forest canopy and undergrowth to get the effects right. If a feature could be included in level editors that allows vegetation to be 'grown' on any type of landscapes by simply tuning parameters associated with the climate and vegetation preferences, a designer's job would be made simpler. Lessons for simulating the natural distribution of vegetation could be obtained from a branch of ecological modelling related to Agent-Based Modelling (ABM) termed Individual Based Ecology (IBE) (Grimm 1999; Grimm and Railsback 2005). There are four criteria in which IBE is differentiated from other less individual-oriented modelling methods: (1) the degree to which the complexity of the individual's life cycle is reflected in the model; (2) whether or not the dynamics of resources used by individuals are explicitly represented; (3) whether real or integer numbers are used to represent the size of a population; and (4) the extent to which variability among individuals of the same age is considered (Uchmanski J and V. 1996; Grimm and Railsback 2005).

The next section explores a vegetation modelling approach (Ch'ng 2009) and seeks to simplify it for the pilot study.

3. Methodology

Game level design and ecological modelling are two very different fields, but lessons could be learned from the latter. The former sought believability and aesthetics while ecological modelling aims to capture a representation of the physical world within a computer model for understanding aspects of it and for proving theories. This section looks at a simplified model for growing vegetation and natural object placements in game level designers.

3.1 *A Simplified Model of Vegetation Growth*

Models in IBE are complex and use relevant environmental factors associated with the habitat of the biological community. These may include sunlight, temperature, altitude, carbon dioxide, and variations of soil types such as soil texture, acidity, depth, etc. Important factors for plant interaction such as competition for space, sunlight, and nutrients are also accounted for. A game design tool does not need such complexities but requires only factors that have high impact on plant growth. Environmental conditions may also be merged into an average. For example, the many different types of soil can be merged into an average condition using a single heightfield as reference.

A survey of plant ecology (Weaver and Clements 1938; Tivy 1993) shows that factors such as sunlight, temperature, moisture, and soil are the most influential factors affecting the growth of a plant. Plant competition on the other hand, is a deciding factor for the survivability of species in the environment and the patterns found on landscapes as suggested by studies in complexity science (Resnick 1994; Camazine, Deneubourg et al. 2001).

3.2 *Modelling the Environment*

In the environment, seasonal changes are not implemented. Time is continuous and is connected with growth, reproduction and competition. The environment has a tropical setting.

Factors	Units
Sunlight	$0 < S < 1$ (default=0.8), $v=5$
Temperature	$0 < T < 1$ (default=25 ° C), $v=0.2$
Moisture	$0 < M < 1$ (default=0.4), $v=0.2$

Table 1. Tropical Settings

The stochastic variations v of each factor is simulated for more interesting results. Soil conditions L are merged into a heightfield of value 0-255 normalised to the range $0.0 > L > 1.0$. Lower values (0.0~0.3) describe better soil conditions, and higher values (0.7~1.0) describe soils that are difficult to grow in.

3.3 *Plant Genotype*

Plant genotypes are defined using the relative measure (Ch'ng 2007). Each allele describes its capacity for adaptation to a certain environmental parameters and its growth and reproduction traits:

Description	**Lower**	**Ideal**	**Upper**
Sunlight	0.3	0.9	1.0
Temperature	-10	38	40
Tolerance to Crowd	0	0.2	0.3
Soil	0	0.4	0.6
Age, Growth and Reproduction			
Energy		1	
Seed Count		5	
Maximum Age		30	
Reproduction Age		3	

Table 2. Plant Genotype

3.4 *Plant Phenotype*

Plant phenotypes are graphical representation of the various stages of growth. Four simple representations are shown in Figure 1. These four stages are seedling, young plant, mature plant, old plant.

Figure 1. Representation of a species of plant in various stages of growth (he representation is a herb)

3.5 *Principles of Plant Growth*

Growth, reproduction, and plant interaction are described using algorithms. Here we define some simple rules for vegetation interaction.

```
FOR EACH PLANT:
  Fitness (f=1.0)
  Energy (e=1.0)
  Increment age every t seconds
  Compute fitness f
    Compute sunlight, temperature, soil
    Interact with surrounding plants
      If f=0, reduce energy loss l
      If e=0, remove plant
  If probability (p<0.5), reproduce offspring
    Reproduction based on fitness:
      Scatter number of seeds (seedCount*f)
        at distance (d) around this plant
END FOR EACH
```

Table 3. Algorithm for growth, reproduction, adaptation, and interaction

The fitness is the deciding factor for plant survivability. The fitness also decides the reproduction capacity of the plant; it contributes to the number of seeds that are reproduced defined in the Seed Count allele (Table 2) and is computed by multiplying it with the seed count: `seedCount * f`. The energy loss l is added for plant resilience to death, this by default is $l=0.1$. Each new seed carries the same genotype as the parent plant and is scattered at a distance surrounding it. Age is incremented by the seconds.

The rules of interaction between a plant and its environment is measured by Ch'ng's Adaptability Measure (*Am*) (Ch'ng 2007). The *Am* measures each environmental factor and generates the simple fitness (Eq. 1) in each system cycle:

$$f = cLUT \quad (1)$$

Where c is the competition, L is the condition of the soil, U is the sunlight, and T is the temperature. Competition in particular occurs if the plant diameter intersects:

$$\sqrt{(O_x - u_x) + (O_y - u_y)} - [O_{size} + u_{size}] < 0 \quad (2)$$

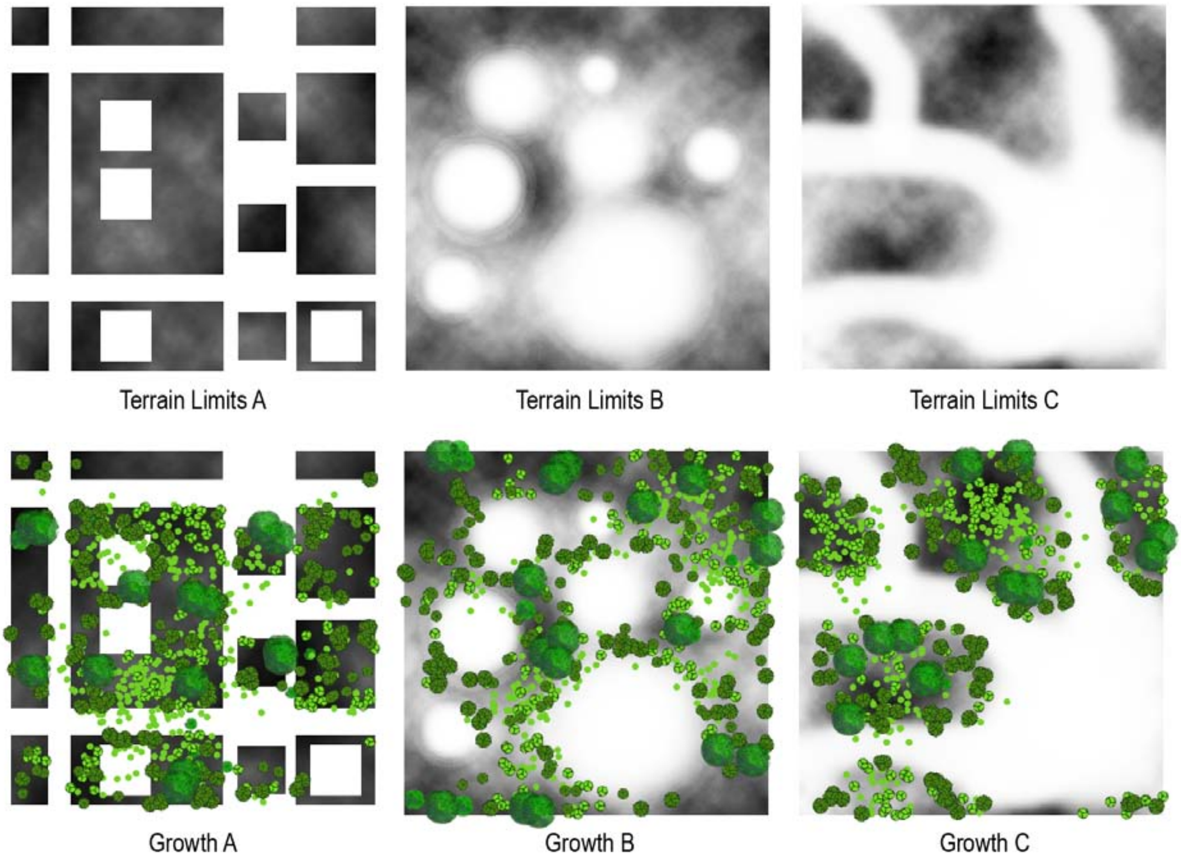


Figure 2. Terrain limits and growth scenarios

Where $O_{x,y}$ refers to the position of the opponent plant and $u_{x,y}$ refers to the position of the plant. O_{size} and u_{size} are respectively the diameter of the competing plants.

A neighbouring plant is a threat if $O_{size} > u_{size}$. The space used by an opponent plant is accumulated only if it is smaller (a herb) than the opponent (E.g., a tree). Each opponent plant contributes 0.1 to the accumulated space *c* and is limited within the range $0.0 < c < 1.0$.

3.6 *Plant Position Conversion: From 2D to 3D*

The position of the plants is within the screen coordinate systems. To convert it to the 3D coordinate system in DirectX, a formula is given below. Conversion to the OpenGL 3D coordinate system is accomplished by flipping the z axis:

$$x_{3d} = \frac{W_{terrain}}{W_{heightmap}} x_{screen} - W_{terrain} \quad (3)$$

$$z_{3d} = \frac{-H_{terrain}}{H_{heightmap}} y_{screen} + H_{terrain} \quad (4)$$

Where $W_{terrain}$ and $H_{terrain}$ are respectively the width and height of the 3D terrain, $W_{heightmap}$ and $H_{heightmap}$ are the width and height of the height map, x_{screen} and y_{screen} are the screen coordinate position of the plant, and x_{3d} and z_{3d} are the 3D coordinate position of the converted plant position.

For ease of export/import to game packages, the information associated with the generated plants in the landscape (plant type, position, and age) are output as XML file (see Table 4). For clarity, the structure shows only five plants. The ‘Position’ tag shows the converted axis.

```xml
</Plants>
  <Plant>
    <Type>Tree</Type>
    <Age>9</Age>
    <Position>34.3472819 52.2839478</Position>
  </Plant>
  <Plant>
    <Type>Tree</Type>
    <Age>15</Age>
    <Position>14.2039487 22.1829918</Position>
  </Plant>
  <Plant>
    <Type>Herb</Type>
    <Age>6</Age>
    <Position>54.2930029 32.1290039</Position>
  </Plant>
  <Plant>
    <Type>Herb</Type>
    <Age>4</Age>
    <Position>61.1920394 27.9938490</Position>
  </Plant>
  <Plant>
    <Type>Herb</Type>
    <Age>1</Age>
    <Position>71.7758392 67.6685930</Position>
  </Plant>
</Plants>
```

Table 4. Plant type, age, and 3D coordinate system position in the XML format

4. Simulation Scenarios

This section demonstrates the potential of the algorithms in three pilot scenarios. The scenarios show that the genotype of vegetation enables them to grow in different niche, mimicking the patterns in natural environments. Figure 2 shows the scenarios. The upper section of the image shows the terrain limits, an average condition of the soil type. The lower section shows the growth patterns that are based on the genotype defined in Table 4 and 5. There are two plant types – Tree and Herb.

Figure 2A is a ‘city block’ scenario, the white areas are limit regions and may represent hard ground (roads, pavements, buildings). The algorithm of the plants naturally prevents them from growing in these regions. The upper part of the terrain shows little plant growth because they have not yet spread to that region.

Figure 2B shows a village scenario. The circular white regions represent inhabitable grounds (from shallow grounds to hard surfaces). It is a noticeable phenomenon that herbs and bushes grow near to the hard surfaces. This was the functioning of the Adaptability Measure. The observed trees grow in deeper soils.

Figure 3B is a scenario where rivers lead into a pond. In this scenario, the gradient (white to black) represents soil humidity. Plants that are adaptable to the such condition grows well.

Description	**Lower**	**Ideal**	**Upper**
Sunlight	0.3	1.0	0.9
Temperature	-10	38	40
Tolerance to Crowd	0.0	0.4	0.6
Soil	0.0	0.2	0.3
Age, Growth and Reproduction			
Energy		1	
Seed Count		5	
Maximum Age		30	
Reproduction Age		3	

Table 4. Tree alleles for four environmental conditions

Description	**Lower**	**Ideal**	**Upper**
Sunlight	0.3	1.0	0.8
Temperature	-10	38	40
Tolerance to Crowd	0.0	0.7	0.9
Soil	0.0	0.6	0.8
Age, Growth and Reproduction			
Energy		1	
Seed Count		5	
Maximum Age		10	
Reproduction Age		2	

Table 5. Tree alleles for four environmental conditions

A culling factor was introduced into each scenario for controlling the population. This is useful as certain game scenarios require dense vegetation whereas other game types may require minimal ground cover. Although restriction applies, it is not strict; plant population may at times grow up to 30% of culling limit. Table 6 shows the culling limits:

Scenario	Population
A	400
B	300
C	500

Table 6. Culling limits for growth scenarios

5. Conclusion and Future Work

This article proposed a novel method for game level designers to efficiently grow natural looking vegetation cover on terrains by simply specifying plant genotypes and environmental limits. The approach, based on ecological modelling, allows designers to specify terrain limits and manipulate the genes of simple vegetation in order to create ground cover for game levels. Three pilot scenarios demonstrated that the algorithms could be integrated into game editors. It could also be developed as an external software package with the exportable XML output of plant types, age, and position. The article also provided a formula for converting from screen coordinates to 3D coordinate system. At present, the software framework is being developed for that purpose. Research is also being conducted to grow vegetation as part of the gameplay for real-time 3D games.

References

Camazine, S., J. L. Deneubourg, et al. (2001). Self-Organization in Biological Systems. Princeton, NJ., Princeton University Press.

Ch'ng, E. (2007). "Modelling the Adaptability of Biological Systems." The Open Cybernetics and Systemics Journal **1**: 13-20.

Ch'ng, E. (2009). An Artificial Life-Based Vegetation Modelling Approach for Biodiversity Research. Nature-Inspired informatics for Intelligent Applications and Knowledge Discovery: Implications in Business, Science and Engineering. . R. Chiong. Hershey, PA., IGI Global.

Gillman, M. and R. Hails (1997). Introduction to Ecological Modelling: Putting Theory into Practice (Ecological Methods and Concepts). Oxford, WileyBlackwell.

Grime, J. P., J. G. Hodgson, et al. (1988). Comparative Plant Ecology: A functional approach to common British species. London, Unwin Hyman Ltd.

Grimm, V. (1999). "Ten years of individual-based modeling in ecology: what have we learned and what could we learn in the future?" Ecological Modelling **56**(1999): 221-224.

Grimm, V. and S. F. Railsback (2005). Individual-based Modeling and Ecology. Princeton, New Jersey, Princeton University Press.

Huber, G. (2004). World Construction Set. Morrison, CO, 3D Nature. **2009**.

Lecky-Thompson, G. W. (2008). AI and Artificial Life in Video Games. Boston: MA, Charles River Media, Course Technology CENGAGE Learning.

McClure, D. (2009). Digital Element World Builder. Oakland, CA, Digital Element, Inc.

Minsky, M. (1980). "Telepresence." Omni **2**(9): 45-51.

Resnick, M. (1994). Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds. Cambridge, Massachusetts, MIT Press.

Reynolds, C. W. (1987). Flocks, Herds, and Schools: A Distributed Behavioral Model. Computer Graphics, Siggraph '87 Conference Proceedings. Anaheim, CA, USA. **21:** 25-34.

Tivy, J. (1993). Biogeography: A Study of Plants in the Ecosphere. Essex, UK, Longman Group Ltd.

Uchmanski J and G. V. (1996). "Individual-based modelling in ecology: what makes the difference?" Trends in Ecology & Evolution **11**: 437-441.

Weaver, J. E. and F. E. Clements (1938). Plant Ecology. New York, NY, McGraw-Hill Book Company, Inc.

BIOGRAPHY

Dr Eugene Ch'ng is currently a Senior Lecturer at the School of Computing and Information Technology, The University of Wolverhampton, UK. He holds a PhD in Electronic, Electrical and Computer Engineering at the University of Birmingham and was a Post-Doctoral Research Fellow at the same department. He has in the past worked on a number of interactive 3D simulation projects related to the reconstruction of terrestrial landscapes (Individual-Based Modelling of biotic and abiotic components), palaeoenvironments and heritage artefacts, the UK's defence arena related to anti-terrorists simulators and training systems, as well as various professional consultations in software, interactive digital media and web applications development. His research areas include Artificial Life, Complex Adaptive Systems, enhanced Virtual Environments (eVE) for Biodiversity Informatics related to marine and terrestrial environments, and Mixed-Reality Interfaces for eVE. The unique strength of Eugene's interdisciplinary field is the fusion of modelling and simulation with interactive 3D visualisation.

An XNA-based Form Control for 3D modeling software

Yindong Yu[1], Rostyslav Lesiv[2], Dr. Rammohan K. Ragade[1]
University of Louisville

Abstract

In the field of game development, Microsoft XNA framework is becoming a more and more popular API. However, due to its effectiveness in setting up DirectX-driven graphical environment, the XNA framework also has great potential when designing graphical-oriented windows applications, such as 3D modeling software. In this paper, we describe how XNA facilitated the design and implementation of a flexible windows form control that allows display and examination of dynamic 3D objects, which is used in an example application to visualize a 3D surface meshing algorithm for finite element analysis.

Keywords: XNA, dynamic 3d model rendering, windowed application, windows form control

Introduction

3D modeling software is a critical element in any computer-aided engineering design. Traditionally those software were designed using C/C++ APIs for DirectX, OpenGL, or even GDI graphics, and therefore generally unable to take advantage of a managed runtime. Also, due to the requirements of writing repetitive boilerplate code, it is often difficult to quickly creating a prototype of a 3D rendering module even for simple uses. The emergence of XNA offered new solutions to address these problems.

Through the design and implementation of a fully-functional windows form control for rendering dynamically-generated 3D models, we intend to show in this paper how such a task can be done with straightforward coding effort using the XNA framework. Because our task required the use of some relatively lower-level APIs in the XNA framework, in the Algorithm section we highlight some of the key API elements we used, and the way we used them. When reading this paper, we assume the readers to have basic concepts of computer graphics, as well as having some simple XNA programming experience. For readers unfamiliar with XNA, we encourage them to refer to the XNA resource websites [1][2] for detailed implementation-related problems.

Data structure

In order to programmatically generate 3D graphical data, a set of basic data structures is needed. XNA already provided plenty of data types for 3D geometry, such as Vector3, Plane, etc. However, due to specific program needs to generate dynamic objects, as well as the possible need for creating a relatively isolated rendering control, it could sometimes be helpful to define custom data types that describes the required graphical information for rendering any object on the screen. These custom data type need to be compatible with the XNA data types so it would be possible to quickly construct the necessary data sent to the graphics device through the XNA API. Based on the above considerations, we designed the following data structure for our rendering data:

RenderObj: a render object is a set of graphical data that would be rendered in a single draw call. It contains geometry information as well as effect parameters for the object.

The geometry inside each *RenderObj* is described by a list of points, a list of indices, and a *PrimitiveType* parameter describing how to draw these points. It is possible to choose among any of 6 primitive types supported by XNA: *LineList*, *LineStrip*, *PointList*, *TriangleFan*, *TriangleList* and *TriangleStrip*. The Color parameter is supplied to determine color for the object. In addition, a line thickness parameter can be supplied to render *LineList* or *LineStrip* with specified thickness.

After the data structure is designed, we also created a helper class called *RenderModelHelper* to generate some basic 3d objects, such as sphere and cylinders. Such a class would help facilitate the generation of surface control points that we need to display later.

The Rendering Algorithm

Graphics Device Setup

One of the biggest advantages of XNA is simplifying the coding effort to initialize the graphics device. For a 3D modeling application that does not use any content pipelining feature to import data, only three steps are required to set-up the rendering environment:

1. Creating a *GrpahicsDevice* object from a set of presentation parameters. Specifically, *IsFullScreen* needs to be turned off when used in a windowed application. Also *EnableAutoDepthStencil* needs to be enabled to process alpha blending. The handle of the form control is passed to the constructor of *GraphicsDevice*, which causes the rendering to show inside the control.
2. Setting up the vertex buffer. In our particular application, we use *GraphicsDevice.DrawUserPrimitives()* to make each draw call, so that no vertex buffer is really necessary. This simplifies the implementation when the 3D object need to be manipulated during rendering.
3. Setting up the effects. An *Effect* describes parameters such as material and lighting. When no material is used in the application, only two effects are needed for rendering lighted and non-lighted objects.

   In addition, the *DeviceReset* event of the *GraphicsDevice* needs to be handled to reset the device parameters when needed. Since no resources were loaded to the graphics memory in our application, no data needs to be reloaded during device reset.

Pre-Processing Render Object Data

In order to optimize for performance, the algorithm gathers all the vertex data into a single array. A struct called *RenderComponent* is created to store the index to the large vertex array for each render object, and the effect used to render it. Whenever a model is passed from the application for rendering, a pre-processing routine is performed to fill the large vertex array with data from the render objects (grouping objects with similar rendering effects together), and create *RenderComponents* to track the array indices. The pre-processing routine also generates non-lighted cylinders for rendering 3D thick lines because XNA does not natively support it. Finally, the desired effect (lighted or non-lighted) are stored together with each render component so that a loop on all *RenderComponents* would provide access to all information required for rendering the 3D model.

The Rendering Loop

Once all data were gathered together, in the overridden *OnPaint()* method of the form control, a *Render()* method is called to update the rendering. The method performs the following operations:

1. Checks if the device is lost (for example, after coming out of screensaver) and recovers it when necessary.
2. Reconfigure *GraphicsDevice.Viewport* according to the updated client width and height.
3. Call *GraphicsDevice.Clear()* to clear the frame.
4. Set up the *World*, *View* and *Projection* matrices for use in all the effects.
5. Render all opaque objects.
6. Disable depth buffer writing and render all transparent objects.
7. Call *GraphicsDevice.Present()* to perform the rendering.

   The procedure used in step 4 and step 5 simply loops through the *RenderComponents* created during the pre-processing stage, and use the index into the vertex array to make *GraphicsDevice.DrawUserPrimitive()* calls. The correct effect for rendering each RenderComponent will be assigned before each draw call.

User Interface

Examination of the Rendered Model

In a 3D modeling application, it is usually desirable to allow the user to examine a 3D model using operations such as rotation and zooming. Such operations could easily be implemented through handling of *MouseUp*, *MouseDown* and *MouseMove* events in the form control, and updating the World and Projection matrices according to user input, following these basic matrix construction equations:

$M_{World} = M_{origin} * M_{rotate}$

$M_{Projection} = M_{orthographic} * M_{zoom} * M_{pan}$

The XNA library provides a powerful Matrix class with methods to easily create Scaling, Translation, Rotation and Orthographic projection matrices, so all of the above matrices could be computed without much coding effort. The only special step needed in our application is to compute appropriate distances for the near Z-plane and far Z-plane used as parameters when constructing the projection matrix. Since the center of the rendered model is always translated to the origin using the origin matrix, computing the maximum distance between two points on the model is sufficient to determine a good distance between the two Z-planes.

Visual Manipulation of Rendering Data

Another common user interface requirement for a modeling application is the ability to select parts and manipulate the model on the screen through clicking and dragging. In our application, we implemented a ray-based selection algorithm which shoots a ray from the point of mouse-click directly into the screen, then gathers points on the rendering model with shortest normal distance to that ray, and sorts them in order. If the ray intersects a 3D triangle, such an algorithm would find the exact coordinate of the intersection point. Therefore, for solid models, any point on the surface of the model can be selected with great accuracy. Also, because the points were discovered through looping through the render objects, once a point is selected, it is possible to return the parent render object containing the point and thus allow user interaction with such object.

Once the selection algorithm is implemented, dragging of object can be easily achieved using event-driven interfaces. In order to update the position of object being dragged in real-time, the form control could be notified with the translations on any render object during each mouse-move step, and apply the translation on the array that holds vertex data. Since *DrawUserPrimitves()* is used for the draw call, the changed vertex data will be automatically reflected in the next frame with no need for updating the vertex buffer.

Application

Interfacing with B-Spline Surface

To demonstrate the capabilities of the rendering control we developed, we generated B-Splinie surface models for finite element meshing and visualization purposes. The surface has four boundaries defined by B-Spline curves, and it may contain an arbitrary number of control points to determine both the curvature on the edges and the curvature on the surface.

Because the parameterized description of a B-Spline surface allows for finding the coordinate of any point on either the boundary edges or on the surface, it is simple to visualize the surface by sampling an arbitrary number of evenly-spaced points to form a grid on the surface, and then triangulate the grid to create render objects. We further took advantage of the user interface features of our form control to allow the user to edit the surface through dragging and adding control points. An example screenshot of surface rendering can be seen in Figure 1.

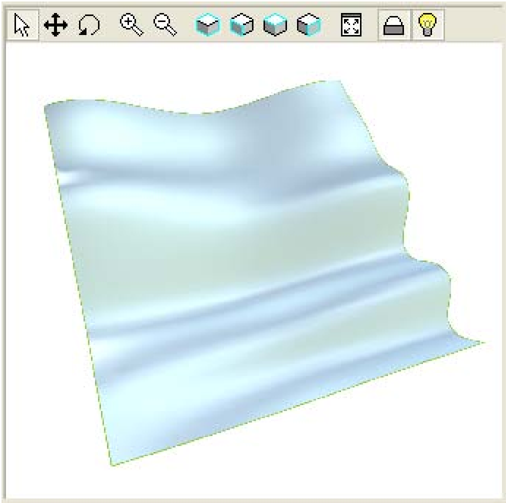


Figure 1 An example surface rendering with lighting effects

Visualizing the Surface Meshing Algorithm

Once the surface is created, a finite-element meshing algorithm is performed on it. At different stages of the algorithm, rendering objects were generated to visualize the progress. The visualization of two intermediate meshing results generated by the algorithm can be seen in Figure 2 and Figure 3.

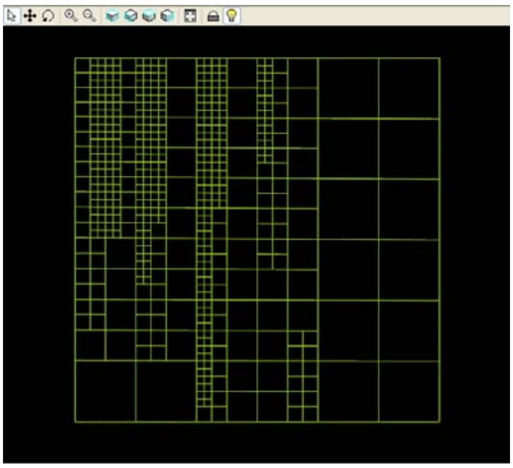


Figure 2 Result from background meshing

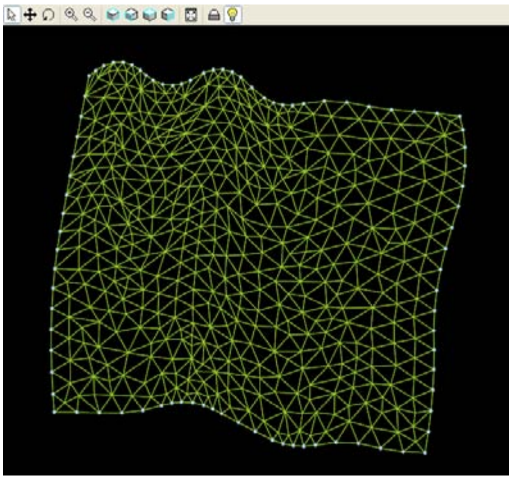


Figure 3 Triangular mesh result from the advancing-front algorithm

Conclusion

The XNA framework turned out to be a brilliant choice in our application, not only because of the beauty of managed code, but the API itself has surprisingly short learning curve. Even when we were not using higher level classes such as the Game class that XNA developers typically start application with, the steps to construct a 3D rendering interface is straightforward and required practically zero specific DirectX or graphics hardware-related knowledge.

The ease of use came at a slight cost – the XNA framework requires your graphics card to have certain capabilities, such as pixel shader 2.0 and vertex shader 1.0. On any modern graphics card though, this is not a problem.

We also note that the existence of *DrawUserPrimitives()* method greatly facilitated rendering of dynamic content, as from the programmer's perspective, you are basically rendering data directly from system memory, without any concern for all the asynchronous data exchange between the main memory and the graphics card memory. XNA even performs the garbage collection automatically so one doesn't even need to worry about cleanup. With the help of that method, it would be easy to realize animation using a timer to change vertex data in real time, as well as changing the rendered object to reflect some dynamic data, e.g. updating vertex colors to display a contour plot from finite element analysis. In an extended version of our rendering control, we have developed both of these features to a great success, and deployed the control inside a real mechanical engineering design application.

References

1. The MSDN XNA Developer Center: http://msdn.microsoft.com/en-us/xna/default.aspx
2. Riemers XNA tutorial: http://www.riemers.net/eng/Tutorials/XNA/Csharp/Series1/Starting_a_project.php

Acknowledgements:

This work is supported in part through a project contract from the US Army Tank-Automotive and Armaments Command (TACOM). It is also based on some additional work done in classwork. We also wish to acknowledge the following team members of this project for their encouragement: Dr. G. Prater, Mr. Robert A. Meyers, as well as Dr. Dar-Jen Chang for issues related to XNA.

Learning Agents in Board Games

N. E. Bassey and Q. Mehdi
Games Simulation and Artificial Intelligence Centre (GSAI)
School of Computing and Information Technology, University of Wolverhampton, UK
Nsikak.Bassey@wlv.ac.uk

Abstract

Artificial Intelligence (AI) in games has become an integral part of game design architecture. A lot has changed since the days of scripted computer games that offered little challenge to human players or were so predictable to be non-immersive. Artificial Intelligent agents have found successful applications in industry, economic and robotics to name a few. In computer games, the intelligent agents known as "Non player characters" (NPCs) are considered to be autonomous entities. Based on scenarios, these agents may be self interested or cooperating to solve a shared problem. The scenario presented in this work is how simplistic learning can be achieved by an agent to exhibit Intelligence using an experimental set up of "TAG DELUXE**"** between our learning Agent Alpha which is pitted against other agents (or human) with no learning. The Application has been coded in the Java programming language with a graphic user interface for visual purposes and input to the application. The Alpha agent shows learning by using a reinforcement learning approach, it is also a reactive agent and shows an abstraction of the potential of these agents to learn from experience or make cognitive decisions. As computer game players demand better human like cognitive capabilities from NPCs, so will the application of intelligent agents try to bridge the gap to provide adequate Artificial Intelligence to simulate human decision processes.

Key Words:
Learning Agents, NPCs, AI, Training, and Reinforcement learning.

Introduction

This paper investigates the use and extent of artificial intelligence in Computer Games with an emphasis on Non-Player characters. A simple example of a learning agent will be demonstrated to show the concept of learning through the development of an application called "Tag Deluxe".

Advances in modern computing have led to vast improvements in game play graphically and visually. While this is good, the Artificial intelligence aspect has been slow to catch up. The traditional AI designs have not had significant impact on the intelligent systems in games. Modern approaches to AI with respect to autonomous software are beginning to be popular (Cressy, 1997). This leads us to think about intelligent programs. Most Authors tend to look at these software programs as agents. These may be computational entities or robots that can be viewed as perceiving and acting upon their environment (Weiss, 1999) which leads to view them as intelligent agents. Developers have tried to use Artificial intelligence to give some form of cognitive capability to Agents (NPCs).These NPC's don't have to be ultra smart or dumb as both types provide a rich variety to game play (Bourg and Seeman, 2004). According to (Priesterjahn , 2007) "Scientific research in the area of game AI has therefore concentrated on the deployment of learning methods to create competitive agents with very high performing results". There are various game AI techniques which are both deterministic and non deterministic. Nowadays graphical advances drive the game industry but in an attempt to increase realism and appeal, the reasoning capabilities of characters "living" inside these virtual worlds must be addressed (Ponsen et al, 2007). Currently in many games, AI still lacks adaptive behavior in the form of self correction (Bakkes and Spronck, 2005). In the case of a predator-prey game, many interesting approaches have been addressed which include Korf's heuristic, co-ordination strategies, multi-agent reinforcement learning (Reverte et al, 2007). These systems are developed for a particular game. Notably components developed for one game can not be easily incorporated into a game produced by a different software developer with the result that code reuse are limited (Lees et al, 2006). In this work a simple approach towards achieving a desired type of Artificial Intelligence is used to design and develop an agent with learning ability.

Artificial Intelligence in games

Generally, there are no standard guidelines for what methods come under AI but one proposed by Wright and Marshall (2000) defines it as "the high-level control code that determines the behavior of game agents". Many games are not intelligent in the human context but merely create the illusion of intelligence. Human intelligence on the other hand is a function of the state of knowledge and what problems they are applied to. In other words it is a domain dependent intelligence i.e. in a specified virtual world; the NPC is able to perform as well as any human or any other NPC (Suliman et al, 2001). A lot of game developers have researched ways to model human behavior in order to produce better AI, by encoding into the AI the actions a human would carry out in various situations. This is to produce AI responses that are humanly realistic (Sweetser at al. 2003). Furthermore NPCs should have the ability to manage memory which includes gathering information about objects in its virtual environment and store copies of the states of these objects and /or other NPCs.

Artificial intelligence in games is becoming increasingly realistic, this is so apparent in their representation using graphics. The problem is trying to improve the believability of the living entities residing in these virtual environments (Ponsen et al, 2007). A lot of development has gone into graphics that have enabled developers to create worlds from cartoon like graphics to near life-like virtual worlds with well modeled characters. These characters look well made but sadly do not match their believability. This has led to a mismatch between the playability and believability of these virtual environments. This current scenario has led to a revival of interest in using traditional academic AI methods such as Genetic Algorithms, Neural Networks and a host of others within Computer games (Delgado-Mata and Ibanez-Martinez, 2007). Most modern game developers require some level of effectiveness without sacrificing speed. Therefore simple and easily understood AI like scripting and finite state machines (FSMs) are still dominant. In regards to board games, an example is algorithms that can extensively analyze future states to use in evaluating outcomes. This is due to their tactical nature and their branching factor is small while in complex modern video games, their game worlds are vast and mostly continuous. This is a problem compared to the defined worlds of Chess (64 positions) or the 361 positions of Go. It is easier for the board game AI to exhaust search options and very difficult to achieve in the same time interval in modern video games for game playing to be smooth or the issue of game agents exhibiting strange or erratic behavior. The other noticeable problem is in board games, action is turn based while in modern video games action is simultaneous (Tan and Chen, 2008). According to Suliman et al (2001) Non-Player Characters can be made to show more intelligent behavior by being able to analyze and make decisions about their environment and other NPCs. This increases the believability and attractiveness of the game. During the game development phase, the game designers can create avenue for the NPC agent to show their believable intelligence. This is somewhat achieved by providing the NPC with a set of tools (AI or Logic engine) to make inferences from. Traditional artificial intelligence has been made simpler by using sets of soft computing tools. Some of these are artificial neural networks, fuzzy logic systems, probabilistic reasoning and evolutionary algorithms.

Modern AI research aims to increase the immersive nature of AI in games by increasing the difficulty of the game to act as an opponent or an assistant to a human player. Game fun according to Hsu et al (2006) has to involve four characteristics: challenge, fantasy, curiosity and control, all these are factors that motivate game players. These are often the guidelines on which most popular computer games have succeeded. One of the most popular online games today "World of Warcraft" by Blizzard (2005) state that their NPCs are AI controlled characters; it is actually a lot of reactive agents and scripted events. The kind of NPC expected behavior is to maximize goal achievement by choosing the most optimal actions with the information accessible to it (Yildirim and Stene, 2008). How do human players of computer games perceive the intelligence of an NPC? One must note that the role of an NPC is not to compute the best way to win against the human player but to complement his actions so the game remains immersive. Nayarek (2004) believes that "Measures such as cheating are absolutely acceptable as long as the suspension of disbelief is retained". AI implementation should be limited as too much autonomy and unpredictability are undesirable. On the other hand Glende (2004) believes that when implementing AI, characteristics significant to humans need to be considered. He goes on to describe the attributes expected from human controlled agent that should be mapped onto NPC behaviour. These are predictability versus unpredictability, creativity in problem solving, personality, intention and acting autonomously, improvising and planning and most importantly learning. He concludes that using modern computer games as a platform for the Turing test still falls short of our goal but by addressing the attributes needed to judge a game character whether human or computer controlled has helped to move research in the right direction.

Many decisions we make in life are sequential in nature and one of the known class of sequential decision making problems are the Markov decision process (MDPs), Based on a Markov assumption that "the current state depends on only a finite history of previous states". This has led to Markov decision with their most important aspect being that the optimal decision in a given state is independent of earlier states decision maker encountered (Van Eck and Van Wezel, 2008). They further go on to describe in the MDP algorithms that are guaranteed to find optimal policies. These algorithms are described as dynamic programming methods whose only faults are that they are unable to deal with problems that have a high number of possible states. One of the solutions proffered involves a new class of algorithms known as reinforcement learning algorithms.

From an intelligent agent perspective, the environment the agent is located in is in a certain state. The agent observes this state and depending on the state, the agent takes an action. The environment responds with a successor state and reinforcement (reward). The agent's task is to learn to take optimal actions or decisions. This may come in the form sacrificing immediate rewards for future rewards, to obtain a cumulative reward in the long or simply to obtain more information about the environment.

Intelligent Agents

Just like biological agents, software agents are designed to detect input from their environment and make decisions in order to achieve their goals.

Agents have been successfully used in various fields but one budding area of research and development is computer games. The need to enhance the believability of games coupled with the falling cost of hardware has led to a renewed interest in agent technology (Lees et al, 2006). With the advent of AI techniques, various researchers have come up with ways to create algorithms and codes that apparently act and play games intelligently. Historically Researchers have separated games into different categories; two of these are the zero-sum and non-zero-sum games. Zero-sum games are characterized by all the information about the game being available from start to finish. Examples of such games include Checkers and Chess which during game play all the pieces are visible so this is regarded as a perfect information games. Non-zero-sum games on the other hand have a hidden component. This means that not all the information about the game is available at any point during game play with known examples being bridge and poker (Franken and Engelbrecht, 2003). One interesting story of using computing power to simulate intelligent behavior is the IBM game playing program Deep Blue. This program made headlines by defeating the world chess champion Gary Kasparov. This was a symbolic victory for AI because playing chess with a then huge state space was no mean feat. Deep Blue was developed by a very formidable team including the US Chess champion Joel Benjamin. The program relied on a combination of brute force and sophisticated knowledge of chess which was supported by a 32-node high performance computer capable of calculating 200million moves per second using an evaluation function (Apte et al, 2000). Deep Blue does not understand chess, it was simple a program using a powerful computer to apply rules to find a move that leads to a better position according to an evaluation criteria programmed by chess experts. Kasparov did not lose to machine intelligence but rather to brute force computation although he expressed doubts that his opponent was a computer and felt more like he was playing a human opponent. This partially agreed with Turing's view in that Deep Blue played chess intelligently but one must also note that the computer understood any of it than a television understands the images it displays (Buttazzo, 2001).

Games AI

Computer games according to Schaeffer (2001) are ideal testing grounds for exploring the capabilities of computational intelligence. Comparing the real world to that of a computer game, a computer game world has fixed rules, interaction is defined and the scope of research or interest is constrained. Real life on the other hand is always subject to change with a limitless scope and interaction occurs in so many and complex ways. One of the early goals of AI was building a program that could defeat the then world chess champion, a process which took over 50 years (Apte et al, 2000 and Buttazzo, 2001). One must commend the champion Gary Kasparov for withstanding the computing power that was DEEP BLUE (machine: 200,000,000 chess positions a second compared to man: 2 a second). In 1950 Claude Shanon published a seminal paper that laid out a framework for building high performance programs, over half a century later, enormous efforts have been carried out in constructing high performance game playing applications or programs. Many game AI programs have been able to best human opponents in traditional games such as Scrabble, Othello and Checkers while improved programs have been able to play more comparable human levels in chess and backgammon (Schaeffer, 2001). On the other hand AI in games is receiving a lot of research attention and most people agree that believability is more important than truth. Thus the goal of AI in games is generally the same as that of the Turing test. In other words, to create a believable intelligence by any means necessary (Livingstone, 2006).

Software Design and Development

The initial thinking was to design a single agent that would exhibit intelligent behavior. Initial concept considerations was a tic tac toe approach. To design an agent that would be able to think. This was abandoned as tic tac toe was grid limited to 9 squares and would have been limited in trying to demonstrate any form of learning. Some thinking outside the box was required and various scenarios for a controlled environment were thought of. The result was the "*Tag Deluxe*" grid which is expandable to an infinite number of squares/grids. The concept was kept simple and focused on showing an apparent learning pattern. A requirement to evaluate this agent led to further development which improved this agent. Thus with each incremental improvement on a pre existing agent, we tested the agent to see how they fared against each other till we maximized the agent potential. There was focus o

The "Tag Deluxe" Application

Tag Deluxe is an experimental application designed to illustrate believable intelligent behavior. An agent (Player Alpha) has been designed to learn to play against an opponent using a simplistic code. Agent Alpha is pitted against any of three other Agents (Bravo- a simple agent, Delta - a more complex agent and player human- a human opponent). The aim of the application is to show that Agent Alpha can learn from observing any of the other Agents and use this learning to score points in the experimental set up. The tangible reward for all the agents is an increased score.

Software Process Model

A software process model is an abstract representation of a software process. Each process model represents a process from a particular perspective so only provides partial information about that process Sommerville (2001, 2007).

The incremental model is a process by which the development is in stages and each stage is a linear sequence. The initial increment addresses the core requirements of the product and is evaluated by the client. Once the initial stage is complete the next increment addresses the modifications made form the evaluation. Every time the project increments a new function is added to the system. The system then can be tested after each increment to test the new function and the whole system. This process helps fulfill the requirements of the client in order to deliver the functions and features specified.

Concept of the Application

Tag Deluxe uses a turn based approach. Both players score points by different ways. The learning Agent Alpha scores a point by predicting/moving to a square or grid the opposing/dodging player will go to. If Agent Alpha correctly predicts the grid/square the opponent will land, he scores a point. On the other hand if the prey player (Bravo, Delta or Human avoids landing in the same grid/square as Agent Alpha, he scores a point.

Both Agents will start at the same square/grid by default. The player with the highest score at the end of the predetermined number of rounds wins. Note this is not the ultimate aim of the application. The aim is to demonstrate the concept that the agent does in fact carry out some perceived learning.

For the purpose of this experiment, Agent Alpha has been designated with the "  " image. He is a smart agent with an ability to learn while the other Agents have been designated with a generic image "**O**". The reason for this being, one of them is selected at the start of the experiment. Each of the opponent Agents selected will behave differently during game play from being simple in the case of Bravo, more complex Delta to a Human Agent which acts/moves at random.

For simplicity player "1" is represented by and Player "2" by O (Bravo, Delta or Human).

Experiment Setting

The following diagrams are used for illustrative purposes as the main game is played using a graphic user interface made in java language. The game is played on a board of nine grids. These are areas of movement and movement can be in any direction. The actual board size can be edited to any numbers but were using nine for illustrative purposes and to reduce the time Agent Alpha will need to learn as this somewhat represents a finite state machine where there are a finite number of states.

The experimental results which were carried out on Alpha, Bravo, Delta agents and human are presented in Appendix A. For standardization reason, all game sessions were run with various run cycles from 10-200.

UML Modeling

According to Flynt and Salem (2004), almost all successful software systems are built from models. A model is a simple representation of something complex. A model makes it possible for those who want to build a system to visualize the system. It's a way to specify what the system is to do and to guide the construction of the system. A model creates a common language, or a common way of understanding. If you put a model in place, your project can move forward with force and momentum that are otherwise beyond your reach. The Unified Modeling Language (UML) provides a set of elements that allow you to model software systems. This will be used to design use case diagram for the game project.

Guidelines of the Application

1. The application is confined to a nine grids layout. This is for simplicity and demonstration of concept.
2. Number of rounds is a minimum of 1 and any other number decided at the start of the game. This is to allow the Alpha agent to carry out some degree of learning for the game to be challenging. The higher the number of rounds, the greater Agent Alpha will learn and play better.
3. Movement is in any direction. This means that any player can move from one grid to any other grid.
4. Each time the Alpha Agent player does not tag (i.e. lands on the same grid) its opponent, the opponent agent scores.
5. The Alpha agent scores when it correctly predicts the same grid as the opponent agent.
6. The prey scores when it successfully lands on a different grid after each game round.

Alpha Agent

This is the learning agent. Its algorithm allows it to learn based on probability and reinforced learning. Agent Alpha starts with no knowledge of its opponent. As the game begins, Agent Alpha will begin to store a library of results. These results are what Agent Alpha will use to calculate the probability of his opponent remaining stationary or moving to a particular grid after being at a previous grid. Also based on probability calculations, the more times an opponent moves to a particular position, the better it will learn and move to that particular position to score. This helps the agent to maximize its success in making the right decision for its next action (i.e. reinforcement learning action). Agent Alpha is also able to retrieve data used to play an opponent if the same opponent plays against it twice in succession. This has the added function of allowing the learning to be quicker.

Bravo Agent

This is a very simple agent. It has no learning experience and is completely automated. It seemingly moves in a regular pattern. Agent Bravo has a simple algorithm that allows moving in an increment of one grid each time. So within a short amount of time, it will be able to move in a recognizable pattern.

Delta Agent

This is a more complex agent. It has a bias (probability function) built into its algorithm that can be altered. A higher bias will create a tendency for it to move to particular grids while a lower bias will make it moving more randomly. Depending on the bias value, Agent Delta provides a bigger challenge to Agent Alpha.

Human Agent

As the name implies, this is an agent that allows a human to act as an opponent to Agent Alpha in the experiment. Human beings are predictable and unpredictable. How a human will fare against Agent Alpha is left to chance. Interaction between human and other agents was established via the point and click device.

Main frame Tag

This is a java file in which all agents can be found. What the tag file does is to create the Graphic user interface, keeps track of the agents as they run within the application, the time between refreshing the GUI, updates the scores and Agent buttons. Basically it acts as a referee. The main parts of the tag frame are the board size implementation, button creation and score string.

Implementation

All agents were designed in the "Java" programming language including the run time environment. The game frame Tag.java as well as the other components (Agent Alpha, Bravo, Delta and Human are easily modifiable using any text editor program. This allows modifying agent behaviour to suit the developer. It can be run on any windows OS with a java run time environment.

Comparing the three experiments

The approach we have taken is based on reinforced learning and learning patterns, Agent Alpha detects whether there is a tendency of the opponent to move to a grid and recognizing this, it uses this pattern to score. The Alpha Agent has been able to learn from all three opponents to an extent, no matter how sophisticated the opponent is, the Agent has learnt enough to make the action. The Agent managed to score a few times against a human player whose movement selection is quiet random.
We have tried to increase the sophistication of opponents to Agent Alpha using three experiments to gauge how well it would perform. Results presented in Appendix A, show Alpha performs better when its opponent becomes less complex. On the average, Agent Alpha showed some form of learning by being able to score in all tests. Overall, we have been able to demonstrate simplistic learning of Agent Alpha by adopting a reinforced learning approach.

Conclusions

In this work, an experimental application was developed to demonstrate how simple learning can be achieved by an agent to exhibit Artificial Intelligence. The learning Agent Alpha was tested against other agents and human which has previously not trained.
The learning agent was used a reinforced learning approach based on predictions and probability. The results were analyzed and graphs were primarily used to show their comparative results. The evaluation was carried out using three agents and an option of a player or human input to test the learning agent. This agent cannot substitute for a human player in terms of a Turing test approach but it does exhibit the required learning depending on the implementation of intelligence process.

Future work may include implementing a neural network and a better GUI design.

References

Apte, C., Morgensten, L. and Hong, S. J (2000) AI at IBM Research. *IEEE Intelligent Systems Nov/Dec 2000.* Pp 51-55.

Bakkes, S. and Sprock, P. (2005), 7th Internatinal Conference on Computer Games, Angouleme, France, ISBN: 0-9549016-2-6.

Blizzard (2005), World of Warcraft, Blizzard. http//:www.worldofwarcraft.com. Last accessed Sept. 2008.

Bourg, D. M and Seemann, G. (2004) AI for game developers. Sebastopol, California: O'Reilly.

Buttazzo, G. (2001) Artificial Consciouness: Utopia or Real Possibility?

Cressy, C. J., (1997) Towards the design and implementation of computer-modeled opponents for use in entertainment and educational software: Some initial considerations, Computers in Human Behaviour, Vol. 13, Issue 2, pp 181-203.
(http://www.sciencedirect.com/science/article/B6VDC-3SWXYK6-5/2/d9e47a34f224407d8082976381433c7f).

Delgado-Mata, C. and Ibanez, J. (2007) AI opponents with personality traits in Uberpong. In: *Proceedings of the 2nd international* *conference on Intelligent Technologies for interactive entertainment*. Cancun, Mexico: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Flynt, J. and Salem, O. (2004), *Software Engineering for Game Developers*. Boston, MA, USA: Thomson Course Technology. pp 75-110.

Franklin, S. and Graesser, A., (1996) 'Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents', Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages, Springer-Verlag.

Franken, N. and Engelbrecht, A. P. (2003), Evolving intelligent game-playing agents. In *Proceedings of the Annual Research Conference of the South African institute of Computer Scientists and information Technologists on Enablement Through Technology.*

Glende, A. (2004) Agent design to pass computer games. In *Proceedings of the 42nd Annual Southeast Regional Conference* (Huntsville, Alabama, April 02 - 03, 2004). ACM-SE 42. ACM, New York, NY, 414-415. DOI= http://doi.acm.org/10.1145/986537.986640

Hsu, S.H., Lee,F. and Wu M. (2006) An Integrated approach to achieving optimal design of computer games. In: *Expert Systems with Applications 31.p14.*

Lees, M, Logan,B. and Theodoropoulos,G.K. (2006) Agents, games and HLA. *Simulation Modelling Practice and Theory*. 14 (6), pp752-753.

Livingstone, D. (2006), Turing's test and believable AI in games. *Comput. Entertain.* 4, 1 (Jan.2006), p6. DOI = http://doi.acm.org/10.1145/1111293.1111303.

Nareyek, A. (2004), AI in Computer Games. *Queue* 1, 10 (Feb. 2004), 58-65. DOI= http://doi.acm.org/10.1145/971564.971593.

Ponsen, M, Spronck,P., Muñoz-Avila,H. and Aha,D.W (2007) Knowledge acquisition for adaptive game AI". In: *Science of Computer Programming*. 67 (1), pp. 59-75.

Priesterjahn, S (2007), Imitation-based evolution of artificial game players. In: *SIGEVOlution*. 2 (4), pp. 2-13.

Reverte,J, Bernabeu,A., Gallego,F. and Llorens,F. (2007), Overview on AI Approaches to the Predator-Prey Problem and their Application to Computer Games. *International Conference on Computer Games: Artificial Intelligence, Animation, Proceedings of CGAMES'2007*. 10 (1), pp114-115.

Russell,S and Norvig,P (1995, 2003) *Artificial Intelligence: A Modern Approach*. 1st and 2nd Ed. New Jersey: Prentice Hall.

Schaeffer, J. (2001) A Gamut of Games. *AI Magazine,* 22(3). http://www.aaai.org/AITopics/assets/PDF/AIMag22-03-003.pdf. Last accessed 20 Nov 2008.

Sommerville, I. (2001, 2007) Software Engineering. 6th and 8th Eds. Harlow: Addison-Wesley.

Suliman, H., Mehdi, Q.H and Gough . (2001) Logic Development for reasoning and Cognitive NPCs.*International Conference on Computer Games: Artificial Intelligence, Animation, Proceedings of CGAMES'2001*. 2 (2), pp35-41.

Sweetser, P., Johnson, D., Sweetser, J., and Wiles, J. (2003), Creating engaging artificial characters for games. In *Proceedings of the Second international Conference on Entertainment Computing* (Pittsburgh, Pennsylvania, May 08 - 10, 2003). ACM International Conference Proceeding Series, vol. 38. Carnegie Mellon University, Pittsburgh, PA, 1-8.

Tan, C.T and Cheng, H (2008) A combined tactical and strategic hierarchical learning framework in multi-agent games". In: *Proceedings of the 2008 ACM SIGGRAPH symposium on Video games*. Los Angeles, California: ACM pp. 115-116

Van Eck, N.J. and Van Weze, M. (2008) Application of reinforcement learning to the game of Othello. In: *Computers & Operations Research 35* (2008) pp1999 – 2001.

Weiss, G (1999). Multiagent systems: a modern approach to distributed artificial intelligence. Cambridge, London: MIT Press.

Wright, I. and Marshall, J. (2000) Egocentric AI processing for computer entertainment: A real-time process manager for games. *Proc. 1st SCS Int. Conf. GAME-ON 2000,* Imperial College, London pp 34-41.

Yildirim, S. and Stene, S. (2008) A survey on the need and use of AI in game agents“, p*roceedings of the 2008 Spring simulation multiconference*. Ottawa, Canada: ACM pp. 124-131.

Appendix A: Experimental Results

Experiment 1: Agent Alpha versus Agent Bravo

After 20 turns, Agent Alpha (learning agent) will have effectively learnt all the moves that Agent Bravo carries out and starts winning consistently. This is a brief representation of the illusion of learning. If the game were played through a GUI, it would seem that the agent was learning. What we are looking at is an abstraction of learning. If Agent Bravo were to be represented by the more complex agent delta or a human player, then the results would be different. Agents can be built to carry out more complex tasks. Figure 1 shows the results of running Agent Alpha vs. Agent Bravo after 200 runs of the game.

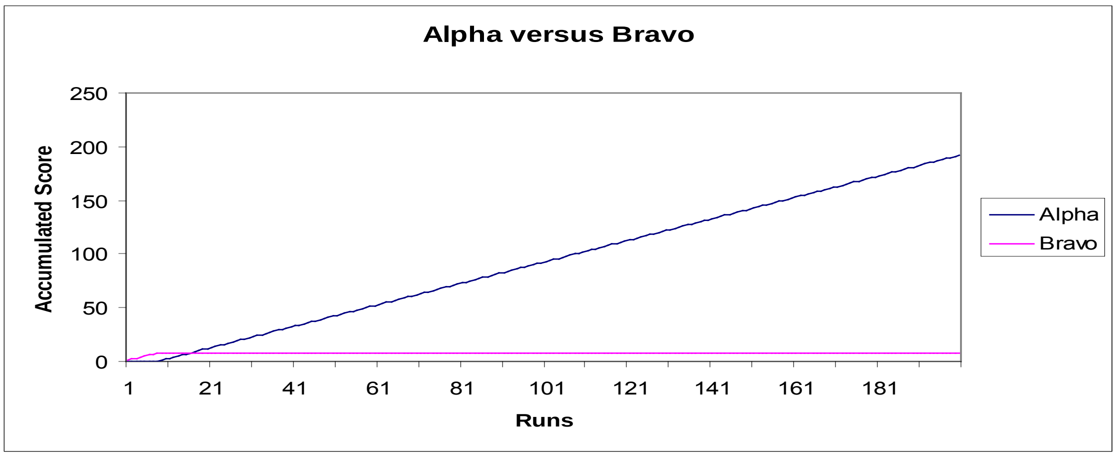


Figure 1: Agent Alpha vs. Agent Bravo

Experiment 2: Agent Alpha versus Agent Delta

Agent Delta is a more complex agent and has a bias algorithm that can be altered by a developer. The bias works by increasing the probability that the agent will go to a particular position. During each game run, it will decide based on probability to return or move to a particular cell. This is what Agent Alpha picks up on and uses to score against Agent Delta. Reducing this bias number will mean that agent Delta will return to a particular number less than often. Increasing the number will increase the probability of Agent Delta returning to a previous position. The results of running Agent Alpha versus Agent Delta with Delat having different bias values (25, 75) are shown in Figures 2 and 3.

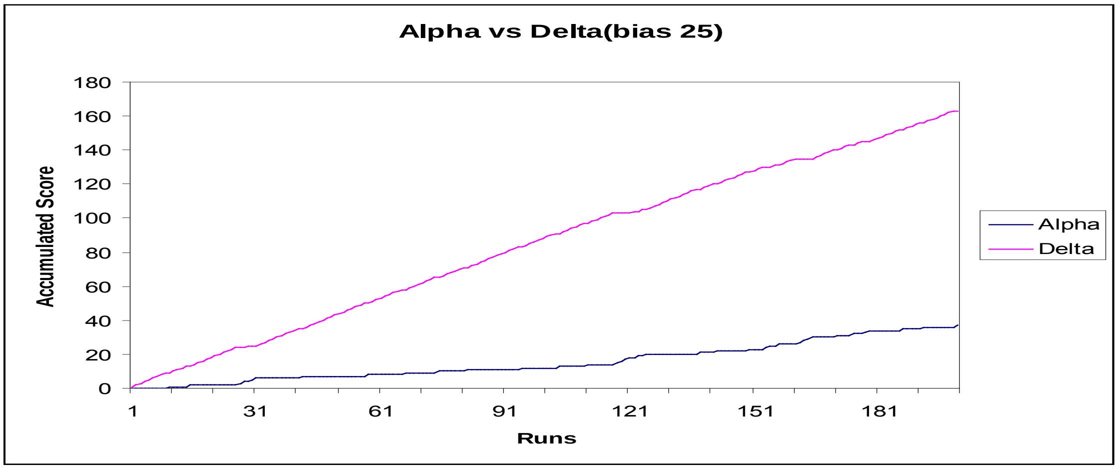


Figure 2: Agent Alpha vs. Agent Delta (bias 25)

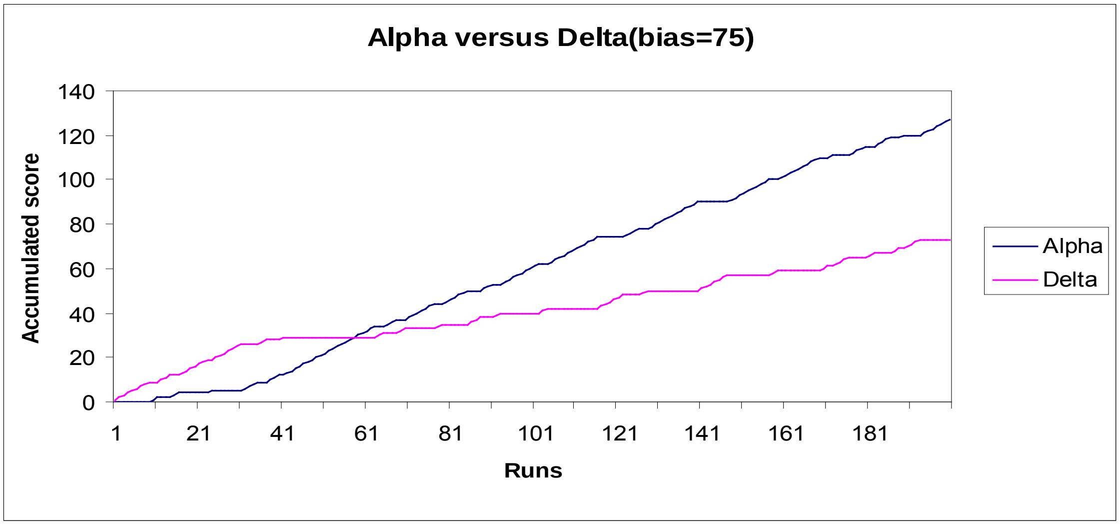


Figure 3: Agent Alpha vs. Agent Delta (bias=75)

Experiment 3: Agent Alpha versus Agent Human

In this experiment, agent Alpha is running against Agent Human (Human input). In this scenario, Alpha's learning ability in limited in comparison with the other agents. However, if human opponent movement is predictive, in the sense that he returned to a particular grid consistently, agent Alpha would have an average prediction and its chances of scoring against a human opponent is high. Overall, agent Alpha does not do very well against a human opponent as a humans movement is too random for any recognizable patterns of movement to be identified as shown in Figure 4.

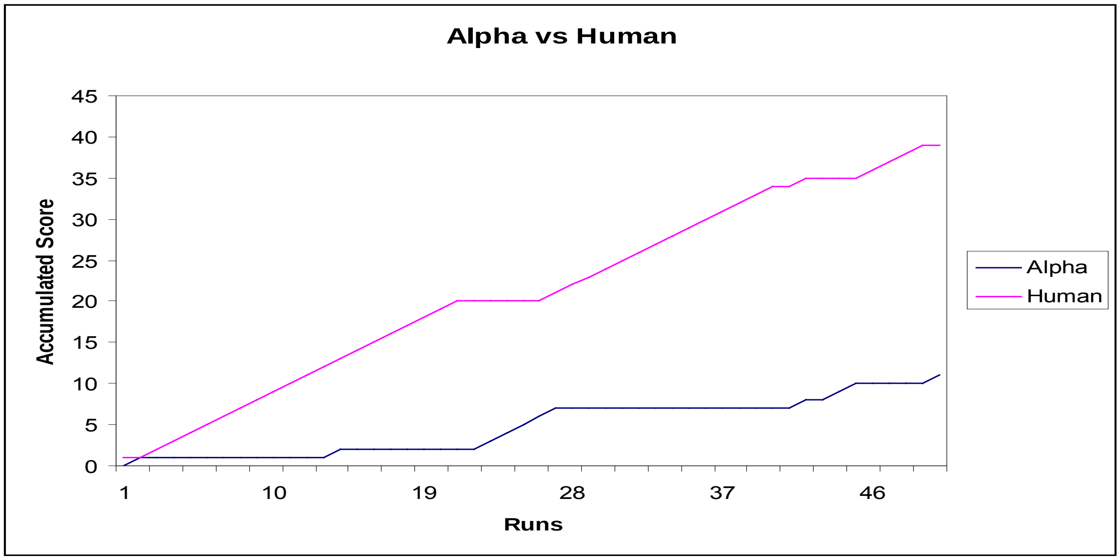


Figure 4: Agent Alpha vs. Agent Human

Authors List