Making the Water Move: Techno-Historic limits in the Game Aesthetics of Myst and Doom

Author Bio
Andrew Hutchison is a lecturer in the Multimedia Design course at Curtin University in Perth, Western Australia, where he has initiated specific game design courses at both undergraduate and postgraduate level. His research is focussed on the relationship between narrative elements and interface in virtual environments, including games and art works. He is one of the producers of the Juvenate project (2001), a multiple award winning new media art work.

Abstract
This paper proposes that the technical limitations at the time of a game’s creation have an enormous impact on the overall aesthetic of any specific game, and also the traditions of the whole craft. Thus, an awareness of this aspect is critical to the useful analysis of games. However, this is often missing from current analysis of games. To illustrate both the significance of techno-historic limits, and several fundamental principals of digital technology, the landmark games Myst (Cyan, 1993) and Doom (id Software, 1993) are explored as examples of the evolution of game aesthetics over time. This leads to an examination of the future limits of the rendering of images and sounds, and how this may impact on future game aesthetics and genres.

Introduction
Because the area of digitally manifested games has developed so rapidly, and the history of games is now significantly long (Järvinen, 2002)(Kent, 2002)(Curran, 2004), there is a real danger that the technical limitations of earlier and current games might be overlooked by some scholars who study games. The issue of the technical limits at the time of a game’s production has been noted by authors in the area (Newman, 2004 p.84), and descriptions of the techno-historic practicalities of gaming systems from a scholarly perspective do exist (Bogost and Montfort, 2007). However, those who are not familiar with the difficult practical aspects of game creation, or who have not studied/had first hand experienced of earlier games, may be unaware of this important aspect. In particular, scholars with backgrounds in more generalised cultural/media studies, film and television, and art, may not have this critical perspective. The rationale for such a detailed understanding of the technical craft history of games has already been made (Montfort, 2006).

Defining “Game Aesthetics”
This paper does not seek to propose a single, simple, resolved scheme for the understanding of a “game aesthetics”. Rather, a definition of “game aesthetics” is presented here only for the purpose of this essay.

Many authors have discussed the uncertain nature of the term “game aesthetics” (Lautern, 2002) (El-Nasr et al, 2006) and have used the term in a variety of ways, often including (or at least not excluding) both the interactive aspects of the players’ experience, and the audio-visual aspects (Klevjer, 2001) (Jenkins and Squire, 2002) (Myers, 2005) (Jenkins, 2005) (Hayward, 2005). Well accepted definitions of “gameplay” exclude the audio-visual rendering of a game, either explicitly (Juul, 2005 p.164), or by not mentioning it (Rouse, 2001 p.18), and tend towards an identification of rules as a defining feature of games (Salen and Zimmerman, 2004 p.158).

Drawing collectively from these ideas, this paper defines “game aesthetics” as the combination of the audio-visual rendering aspects and gameplay and narrative/fictional aspects of a game experience. This linking of the “audio visual” (and other sensory) rendering with the “gameplay” explicitly highlights the important co-dependence that this essay proposes.
Historical Context

Two of the most often referred to computer games are *Myst* (Aarseth, 1997) (Darley, 2000) (DeMaria and Wilson, 2002) (Wardrip-Fruin and Harrigan, 2004) and *Doom* (Manovich, 2001, p.244) (Wolf, 2003 p.62), released within months of each other in 1993.

The *Myst* and *Doom* titles, both being born early in the 1990s and continuing as distinct intellectual properties to the present day, represent an opportunity to examine the evolving technological limits of games aesthetics generally. While digital games certainly existed before *Myst* and *Doom*, the time span covered by these two series includes the remarkable period we can refer to as the “digital nineties”, when digital devices, and especially the personal computer, became an everyday part of the lives of most people living in the modernised world. Both *Myst* and *Doom* were dependant on the personal computer, which not only became commonplace in the 1990s, but also quickly became capable of networking, reproducing colour, sound, animation and video from CD, and then Internet, DVD and broadband video delivery.

Part 1 of this essay describes the many similarities between *Myst* and *Doom*, and how their very different game aesthetics were the result of unavoidable technical compromises, rather than the intentions of the designers. Part 2 examines the future limits of the rendering of images and sounds, and how this may impact on future game aesthetics and genres.

**Part 1: the Myst and Doom Phenomenon**

*Myst*

In *Myst*, fusions of architectural styles and technology dot unpopulated island locations, surrounded by empty seas. A series of puzzles must be solved in order to travel to other island worlds, to recover missing pages of magic books. These can then be used to discover which of three characters in the story is the villain.


![Figure 1. “On the dock” in Myst (1993). The starting view from the first node of the game. The spots in the grass and marble in the left hand side of the image are “dithering”.](image)

*Myst* was developed by the company Cyan (now Cyan Worlds), who had previously produced children’s software, the first such title being *The Manhole* in 1987 (DeMaria and Wilson, 2002, p.258). *Myst* initially ran only on Apple computers, and used an Apple proprietary authoring system named Hypercard. Released in
October 1993, it quickly became a “killer app” for Macintosh computers, prompting users to upgrade, or make their first purchase (Lunenfeld, 2000) (Seward, 2003).

Myst for Microsoft computers was released in February 1994. Then unrestricted by the relatively small size of the Macintosh market, Myst went on to sell around seven million copies (Takahashi, 2002), probably making it the largest selling personal computer game (as opposed to dedicated console gaming systems like Sega and Nintendo corporation’s) till at least the late 1990s (Caroll, 1997). Thus, it is possible to suggest that Myst may be the first widely experienced computer mediated, audio–visually rich virtual reality experience.

Myst presented a constructed, consistently behaving, deep, media rich world that unfolded in people’s living rooms or work cubicles. It was non-trivial, competent both creatively and technically, and it created a sense of “being there” (Laura Miller, 1997). In some ways, it achieved the immersion constantly promised by “virtual reality” hype of the late 1980s (Sandin et al, 2001, p.266). Even those who don’t admire Myst acknowledge its impact on the field of computer gaming, significant aesthetic appeal, and technical achievements (Cyril, 2003).

Myst was not the first CD-ROM delivered game, the first visually rendered adventure game, the first 1st person perspective game, or the first three dimensionally modelled and rendered game. However, Myst clearly was the first time that many people had encountered this kind of interactive virtual experience, or saw this kind of multimedia, and so Myst became a benchmark.

Doom

While 1993 was the year that the Myst/CD-ROM phenomenon started, it was also the year that id Software(sic) released the definitive “first person shooter” Doom (Manovich, 2001, p.244). Like Myst, Doom is also credited with many technical firsts, became a huge popular hit, defined a genre, and has also survived to the present day as a distinct intellectual property. Doom demonstrated the emerging power of two more new media techniques - real time three dimensional rendering, and real time networking. This allowed multiple players to play in the same virtual environment from different computers.

In the same way that Cyan had plenty of previous experience to prepare them for producing Myst, id Software had released Hovertank 3D in 1991 with the claim that it was the first ever personal computer game with 3D elements (id Software, 2005). This claim has been contested, and the game Elite (Braben Bell,1984) has been presented as an earlier example (Wikipedia, 2005). In 1992, id Software produced Wolfenstein 3D, a first person shooter with overall very similar game scheme and technical implementation as Doom. It should be noted though, that there are many other claims for “first networked, three dimensionally modelled and rendered, first person shooter” game – MazeWars (Various authors - 1973 to 1987) (Thompson, 2004) (Montfort, 2003, p.80) and Spasim (Bowery, 1974) (Bowery, 2001).

Nevertheless, like Myst, Doom was a “killer app”, prompting people to upgrade hardware, or even make their first ever computer purchase. Downloaded from the pre World Wide Web internet by technically savy computer enthusiasts, or bought as "shareware", Doom was then passed hand to hand on multiple floppy disks. Many fans bought both the commercial release of the game and a better computer to experience it.

Myst vs Doom

Doom was similar to Myst in fundamental ways. It was also experienced from a first person perspective, and also created from a three dimensionally modelled and rendered world. However, Doom’s game-play was totally different from Myst’s, since it was about killing science-fiction monsters in corridors with a variety of weapons, with no narrative offered (other than the “US Marines in Space” theme borrowed from the 1986 movie Aliens and a brief framing comment in the “readme.txt file”). It utilised real time three dimensional rendering for getting its graphics to the screen, producing a very low resolution, fast moving aesthetic in which sound was not critical. Myst, by comparison, was simply unplayable without sound, not only for the immersive “atmospheric” effects, but in the communication of information critical to the gameplay and narrative.

Myst and Doom quickly came iconic of two completely opposed game forms, with regard to their content, gameplay, over all audio-visual aesthetic and demographic of players (Wolf and Perron, 2003, p.6). This is ironic given that, had the design teams had their way, Doom and Myst would have been audio-visually very, very similar indeed.

The Pragmatic Expression of Myst and Doom

If the technology of 1993 had been no limitation, both Myst and Doom would have had sensational cinema style, photo-real graphics and audio, with the additional feature of interactivity.
The cause of the difference between the audio-visual aesthetic of Myst and Doom is a pragmatic one, imposed by a limitation in technology. Doom did not have cartoon style characters due to a lack of imagination or technical skill on the part of the designers at id Software. Myst did not have still images because Cyan’s designers were limited in their vision and ambition. Both companies solved the enormous technical limitations of the early 1990s in order to make their virtual worlds manifest as best was possible at the time (Halifax, 2002) (Miller, 2002).

The worlds of both Myst and Doom are virtual, mathematical descriptions of the shapes and dimensions of spaces and objects, and their relative positions. A virtual “camera” defines a view of the world, and in both Myst and Doom, this is from a “first person” point of view, simulating that of a normal sized human standing up on the horizontal plane of the virtual world.

The problem was, and is, that to produce the image “seen” by the virtual camera requires mathematical calculation from the descriptions. The more realistic the shape of the objects, the more they overlap, the more detailed the textures and lighting, the more they move, then the longer these calculation will take for each view. In 1993, there were no computers in general use that were capable of doing such large amounts of calculations in a practical amount of time.

Current (2007) levels of personal computers are massively more capable. The central processing units/motherboards of the computers are many times faster. Industry standards for graphics and audio generation now exist. The sophistication of the software programming of games engines has also taken massive steps forward. The effective result is that an up to date computer with suitable video and sound cards is hundreds of times more capable of the real time rendering of images and sounds than in 1993. A good illustration of the increase in computer capacity between 1993 and 2008 (and the resulting sophistication of the audio-visual rendering) is that the original demo for Doom is only 2.3 Megabyte. The 2004 remake of Doom, titled Doom 3, has a downloadable demo of 460 Megabytes. Another measure of the difference is that Doom contained only 54,000 lines of code, while Doom 3 has 785,000 (Kent, 2004).

Thus, we can see that in 1993 personal computers were hundreds of times less capable of creating the images and sounds of a virtual world. Having both high image/sound resolution and real time responsiveness was impossible. Designers were left with the difficult decision of what aspect of their possible environments they would like to implement, and which to sacrifice. Myst and Doom, with different immersive/gameplay ambitions, went in two opposite directions. Myst went the visual “high and slow” road, and Doom went the “low and fast” road.

The Low and Fast Road – Low Resolution Real Time Rendering

Doom’s producers, id Software, needed to create the experience of the user battling monsters, and to make this compelling, they needed to fulfill two prime characteristics. First, to allow “free movement” of the player’s point of view. Second, for this view to respond immediately to the player’s input, as they pressed the key board arrow keys for forward, back, look left and look right. This could only be achieved by “real time three dimensional modelling and rendering” (also referred to as “on-line rendering”). In “real time 3D”, the mathematical descriptions of the world are loaded onto the player’s computer, and the computer renders the images to the screen in direct response to the user’s input. This is happening simultaneously, in real time, with the user’s
gameplaying experience. To achieve this on the common personal computers of 1993, the quality of the image had to be reduced to an almost abstractly low level. The objects and characters were extremely crude, with minimal, repetitive textures. Even these compromises were not enough, and further short cuts had to be taken. *Doom*’s images are a combination of real time 3D rendered techniques *and* the inclusion of more conventional two dimensional animating, pre-rendered graphics. Thus, although it was a technical triumph of its time, the original *Doom* now looks like an abstract art cartoon.

The original playable shareware “demos” of *Doom* still work on current (2007) Windows and Macintosh operating systems, and are available at:

Windows system - http://www.dosgamesarchive.com/download/game/7


**The High and Slow Road - Nodes and Video Files**

Overcoming *exactly the same* technical limitation, but prioritising a different game-play ideal, the designers of *Myst* took the opposite approach to achieve their rich-media “environmental” immersion. They achieved the highest possible level of detail in the images and sounds by sacrificing the features of “free look”, and “immediate response”. The creators of the game used a virtual camera to create four wide angle pictures (north, east, south and west) from many predetermined view points or “nodes” in their virtual world. Because these still images are *not* being generated in real time on the player’s computer (just loaded from CD-ROM), they can be of much higher quality. The resulting visual experience is a series of “landscape paintings” or “postcards” linked together, with occasional live action video clips of actors performing.

The disadvantage of this system is that the player cannot have a view in the world, or go anywhere, that is not provided by the designer of the game. You, as the player, literally cannot take a step wrong. Further, the player’s movement is not a continuous view, since the first person view cuts, or at best, dissolves from one predetermined node to another as “you” move around. The price of high resolution images in 1993 was stillness and discontinuity.

Another consequence of the “nodes” technique is that there is very little redundancy of images/places (due to the time consuming production process). This has an obvious impact on the gameplay - if you go somewhere, it almost certainly indicates that there is a clue there to be seen or heard. This lack of redundancy increases the sense of lack of free movement, and the feeling of stillness.

Like id Software’s *Doom* sacrifices, even when these compromises had been made, it was still not enough. To reduce load time and data size, *Myst*’s images are even smaller than the small screen sizes of the time, and images are often marred by conspicuous “dithering” of colours to give the impression of more resolution (Watson).

However, in what is possibly *Myst*’s greatest aesthetic achievement, the creators of *Myst* utilised the original designed intention of CD technolgy (to stream audio files) to overcome this visual stillness (Fargo, 2003) (Miller, 2002). The visual gaps are filled in with a continuous sound scape eg: the sound of the wind, the lapping of water on the shore. Sound effects are often used to “render” the movement of walls, doors and machinery, so that less animation is required. Music is used sparingly during actual gameplay, mostly to herald the discovery of an entrance to a new place.
Another critical technique for overcoming the stillness of the image in Myst is the movement of specific objects, such as a door opening, the turning of a handle, or the flight of a butterfly across the path. This was achieved by creating an animation sequence of only the relevant part of the screen and playing it on top of the background. Once again, this reduced the amount of data and machine capability required to create the sense of motion in the seen world. The success of this technique was dependant upon the “on-top” animation’s image quality being identical to the background image. Unfortunately, this was not always possible, due to the loss of image quality inevitable in the codecs (compression/decompression algorithms) of the animation files.

Another side effect of creating all of the hundreds of images before the game is delivered to the player’s computer is that a large amount of data is produced, and this is why Myst was dependant upon CD-ROM technology to physically deliver it to the player’s computer - and stream it from CD-ROM during gameplay, since there were no hard drives big enough to store it. Unfortunately, early CD-ROM drives were very slow at delivering files, and this produced quite a slow response and/or “jumping” when the player moved from one node to the next, or an “on-top” animation played. In contrast, Doom’s mathematical descriptions and low resolution textures were originally small enough to be delivered to the computer on a series of floppy discs, transferred to hard drive, and the experience calculated from those small files.
The End of the 1990s
While Myst and Doom are iconic of the 1990s PC games boom, and both have spawned sequels that extend to the present day, their fortunes have been very different. id Software have gone from strength to strength, each new title (including the Quake (1996) first person shooter game series) pushing forward audio-visual rendering in games and consolidating popularity. The Myst series, however, has undergone a dramatic drop in popularity, with sales of each successive title never achieving anything like the original Myst’s success (Myst5, 2004) (Takahashi, 2002) (Hamilton, 2004). Some industry observers characterised Myst’s non violent, story based puzzle solving adventures as “an antiquated style of gaming” (Kasavin, 2004).

At the same time, by the late 1990s, it was clear that the massive steps forward in computer technologies and techniques meant that pragmatic choices between “the high road” and “the low road” would no longer have to be made.

Part 2: The Future Limits of Game Aesthetics
The advances in the audio-visual rendering of games are taken for granted by consumers, and it seems, the games industry itself. However, the limits of these purely audio-visual improvements may not be far away. If the image and sound rendering of games reach a plateau, what will become of the other component of this essay’s loose definition of “game aesthetics”, the gameplay and narrative/fictional aspects. What technical and craft limitations and opportunities might present themselves?

Making the Water Move
A convenient indicator of the development of computer speed and aesthetic ambition of games designers during the 1990s is the treatment of water in the Myst series. By the time of realMyst (2000), the water was fully real time rendered. The surface of the oceans moved and had surface effects like foaming, and the weather changed as you watched. In 1997, Riven had “photo-realistic” still images of the oceans, and managed in one scene to animate the part of the ocean closest to the shore (the “whark cove”). The oceans of the original Myst in 1993 are completely flat and frozen, without any visual motion on them at all, and look quite ridiculous now, compared to any of the subsequent versions.

However it is very important to note that they did not “look” frozen in 1993. Not only was it not technically feasible to do the water effects in 1993, it was not expected. At the time, Myst was the most audio visually detailed virtual experience ever had on a personal computer. In the same way, Doom’s sheer speed of response made it compelling, despite its low resolution. Their aesthetic and immersive impact cannot be denied, as surveyed from the sales figures, and the press of the time.

The success of the earliest, crudest forms of Myst and Doom reveals the truth that all forms of audio-visual rendering, including text and photographic images, depend on a sensationally sophisticated technique of finally smoothing out all of their technical limitations. That technique is, of course, the individual human’s capacity to fully rendering it in their own imaginations (Miller, 1997). However, a further human characteristic, and a fundamental of computer science, work together to quickly defeat this imaginative tolerance of low resolution audio-visuals.
The Moore’s Law Guarantee
In 1965, Gordon Moore predicted that the cost of computer hardware would continue to go down very quickly, even as its capacity increased rapidly (Moore, 1965). This phenomenon, which has held true for forty years and allowed the digital revolution, has become known as Moore’s Law (Intel, 2005).

Game designers, especially of the Myst and Doom series, are in the habit of pushing current technology to the limits. They often produce successive, more sophisticated game experiences that can not work even on quite recent computers, gambling that users will upgrade (Daigle). And users do upgrade. Literally hundreds of thousands of people planned significant computer/video card upgrades around the release of Doom 3 and Half-Life 2 in 2004 (Morris, 2004). What makes this gamble worth taking for the designers and worth the cost for the players?
In 1997, Aarseth specifically accounted for the demise of the text based interactive adventure game thus:

Images, especially moving images, are more powerful representations of spatial relationships than texts, and therefore this migration from text to graphics is natural and inevitable (Aarseth, 1997 p.102).

It seems logical to presume that the same migration is occurring from lower quality graphics to higher quality, driven by this “natural” desire. Every step forward in gaming audio visual sophistication has been met with comments such as “it was like being there”, including Space Wars (Russell, 1961)(DeMaria and Wilson, 2002, p.13), and even the text only Adventure (Crowther and Woods, 1972 -1976). However, this “more powerful representation” only remains more powerful until the human eye sees something more sophisticatedly rendered (Darley, 2000, p.28). The games Myst and realMyst demonstrate this effect particularly well, since the two titles are actually the same gameplay/narrative/world:

Steven Ogden has described the difference between the original Myst and realMyst to be ‘as varied and pronounced as the difference between looking at postcards from a place and actually going there in person.’ (Jong, 2001, para 4)

The 1993 version was a breakthrough, but by 2000, a new cycle of expectation had made it “antiquated”. Each step forward quickly becomes a new minimum standard, and is itself sure to be outdated by the next step forward, which occurs incrementally in desktop PC hardware, and in giant leaps with console hardware. The accolades Doom 3 (id Software, 2004) and Half-Life 2 (Valve, 2004) received in 2004 for their textures, lighting effects and dynamic elements were also awarded to their predecessors only a few years earlier. For example, in 2004, Newman described the original Xbox and PlayStation 2 as producing “near-photo-realistic 3D graphical environments” (Newman, 2004, p.163). In light of the audio-visual steps forward made by the current generation (Xbox 360 and PlayStation 3), it is clear that “near photo-realistic” has a highly transitory meaning. However, each step forward is compelling for end users – however briefly.

For example, gameplay in TombRaider (Core Design, 1996) is very much the same in the PlayStation 3 version as it was on PlayStation 2, and as it was on PlayStation 1, but the game has still sold very well on the basis of the audio visual improvements. The current generation of games consoles (PlayStation 3, Microsoft Xbox 360 and Nintendo Wii) have each sold tens of millions of units in their first year of release (Enderle, 2005).

The Moore’s Law effect has guaranteed an escalating cycle of increasing capability and affordability, commercial competition and cultural/consumer expectation. Both the producers and the consumers of games expect each version to be audio visually bigger, richer, smoother, etc (Magal in Valve, 2004 p.63).

The Spectacular
Perhaps lending strength to this phenomenon is a current popular culture stylistic trend which emphasises “spectacle” over narrative and gameplay. Peter Lunenfeld has identified this broad movement in popular culture generally:

…our culture has evacuated narrative from large swaths of mass media. Pornography, video games, and the dominant effects-driven, high concept Hollywood specturals are all essentially narrative-free: a succession of money shots, twitch reflex action, and visceral thrills strung together in time without ever being unified by classic story structure (Lunenfeld, 2000, p.141).

Andrew Darley says…“Computer games are part of a cultural space of surface play and neo-spectacle…”, adding them to theme parks and music video clips as examples of a new modern aesthetic dominated by sensual stimulation and demonstrations of technical mastery of digital visual effects, in which narrative is displaced (Darley, 2000, p.149).

Aspects of Darley’s analysis of Quake and Myst are problematic and superficial, and in some parts factually inaccurate (eg: his publication date for Myst). Further, like some other cultural studies commentators, he seems...
to place a negative value judgement on games compared to movies or literature. However, his point can easily be applied to games that have shown a marked increase in audio-visual rendering, but little in the other aspects of this essay’s definition of “game aesthetics”. Leading game designers warn about the escalation in graphics, at the cost of other features such as gameplay, and story (Feldman, 2004).

Placing games inside a clustering of music video clips, thin plotted action movies and theme parks would seem to confirm Aarseth’s 1997 concerns about the loss of the opportunity that could have (and might still) arise from the unique structure of interactive adventure games (Aarseth, 1997 p.128).

How Far Can Audio-Visuals Develop?
It is possible that Moore’s Law will come to an end. Certainly, there is no real technological limit to the effective audio-visual improvements that could be made. However, Moore’s Law is not just a prediction about the rate of computer power increase. It is that the capacity will increase at the same time as the dollar cost will decrease. It is this counter intuitive relationship that has allowed for the continual turn over of gaming hardware in the user’s hands, and led to the pre-eminence of Darley’s “spectacle” over narrative and gameplay.

While some speculate that the limits of digital technology are near, Moore’s Law has already surpassed several earlier, predicted “walls” (Twist, 2005). Despite the specific limitations of the current generation of technology, there is plenty of evidence that the improvements will go on (Kanellos, 2005). Projecting the trend of Moore’s Law into the future, we can see that there is very unlikely to be a limit to the improvements in richness, depth, smoothness and fidelity of audio-visually rendered images anytime soon. Ironically, a plateau in audio-visual rendering may soon occur, because the “photo-realism” is actually achieved, with attempts to achieve photo-realism in games “in their final stage” according to some observers (Jenkins, 2007 p1).

However, John Carmack, the famed game engine designer and co-founder of id Software, has suggested that there will still be improvement to be made to the realism of gameworlds in 20 years (Back Door, 2006). He indicates a distinction between pure photo-realism in a non-interactive animation and “realism” in an interactive experience, where the unpredictability of the user’s actions provide many more expectations and possibilities for the rendering engine that mere “photo-realism”.

In 2007, Aarseth’s 1997 observations still ring true:

It is a paradox that, despite the lavish and quite expensive graphics of these productions, the player’s creative options are still as primitive as they were in 1976 (Aarseth, 1997, p.103).

Aarseth was making this point in direct reference to both Myst (“unpopulated and boring”) and Doom (“completely robotic villains”), noting that such problems had been identified in an even earlier generation of games by Mary Ann Buckles in her 1985 PhD study of the interactive fiction work Adventure (1972-1976).

So why does an area as apparently full of potential as digitally enabled games rely so heavily on simple increases in audio-visual rendering and movie style spectacular effects? While there are noteworthy attempts, games are still conspicuously without the capacity to create characters that are dynamic in response to the user’s individuality, or have storylines that are not pre-scripted and limited (Crawford, 2003, p.257).

The Lack of AI, and Abundance of RI

Unfortunately, I think we have gone through a bit of a dark ages as far as communications AI is concerned, but we’ll hopefully come out of that soon (Darling, quoted in Poole, 2000 p.107).

The as yet unfulfilled promise of AI (artificial intelligence) is a conspicuous character of the development of computer games. This limitation is probably the real reason that Cyan Worlds took its Myst related world Uru (2003) title on-line and real time rendered. To go beyond the fixed (though multiple-ending), character poor games of Myst and Riven, Cyan Worlds had to over come the lack of AI with what we must obviously call RI (real intelligence). By bringing players together (on-line) so they could interact, they collectively generate dramatically credible and experientially satisfying gameplay for each individual player. Cyan Worlds did not use
real time rendering in *Uru* just to achieve the feel of free movement and free look, or a current audio-visual “photo-real” aesthetic. It was also necessary to allow the inherent variability needed for player to player engagement, and the dynamic environments they would populate and effect. And *Doom*, while it has always had a single player option, was network multiplayer capable in its first release.

Of course, the value of this use of RI in games has been convincingly demonstrated by the success of several MMORPG (Massively Multiplayer On-line Role Playing Game) titles, of which *Everquest* (Verant, 1999), *Ultima On-line* (Origin, 1997) and *World of Warcraft* (Blizzard, 2004) are now classic examples. It is important to recognise that these games have a clear lineage back through text only on-line MUDs (multi-user dungeon), and then back to non digitally mediated, face to face role playing games, where RI systems were/are the norm eg: *Dungeons & Dragons* (Gygax and Andersen, 1974).

The time when AI might be able to replace RI seems a long way off. In 1950, Alan Turing proposed the “Imitation Game” as a benchmark for artificial intelligence (Turing, 1950). Variations of Turing’s original proposal have become known as the Turing Test, in which, during a text to text conversation, an automated computer programme tries to convince a human that the programme is actually human (Saygin, et al, 2000). The annual Loebner Prize for passing the Turing Test has a one hundred thousand dollar prize, but is still un-won in sixteen years (Loebner, 2005). While “chatter-bots” are already in use in certain limited applications, it is quite dispiriting to have an on-line conversation with any of these recent winners, and find the conversation turn evasive or ridiculous in only three sentences eg:

http://www.jabberwacky.com/chat-joan

The application of typed speech recognition in games, called “language parsing”, is a major feature of interactive fiction (Montfort, 2003), a strong popular form of the 1980s. Aarseth noted that interactive fiction characters of the 1970s and 1980s were “autistic” (Aarseth, 1997, p.115), often ignoring the player, replying with nonsense, or issuing repetitive conversation ending phrases. And this problem has not improved since then. Mainstream games with massive resources behind them still feature characters that behave this way.

The solution to this in *Everquest* (an MMORPG) is to simply have the non-playing characters largely ignore the player after a few lines. In fact, the user’s guide for *Everquest* explicitly warns the player just how limited the language parser is, and explains how to enter only key words in the “conversation” to save time, since the parser software cannot actually understand the whole sentence. In trade exchanges with the non-playing character merchants, a visual menu is provided, rather than risk even relatively simple exchanges to language parsing dependant “chat” (*Everquest*, 2002). In the case of *Half-Life 2* (2004) or *Doom 3* (2004), clicking on non-playing characters to get more conversation will cause them to them repeat variations of three or four conversation ending lines. The *Myst* games avoided these “autistic” conversations by videoing actual actors delivering the lines. However, this just compromises the player’s character instead. The video taped characters just give and take items from you as though you are stupid, mute and physically disabled.

Each year, the Loebner Prize awards a prize for best attempt, as opposed to a successful attempt. It is notable that nine of the sixteen years of the Loebner “best attempt” prize have been won by just three people. This might suggest that there is actually little work being done in the area, as well as indicating how difficult it is. Steven Poole has suggested the whole area of language parsing for games was essential abandoned in the 1980’s, partly because of the rush towards graphics, but also because in the days before personal computers were common, there were no keyboards that any language could be typed into, since console systems did not have them (Poole, 2000 p.106). The whole area of natural language recognition by computers is regarded as important to the future development of games, but also considered very problematic, with no solution in sight (Mateas, 2004). Hugh Loebner, the founder of the Loebner prize, does not expect it to be won in his lifetime (Graham-Rowe, 2006).

Mimesis Engines
There is no shortage of commentary about the limitations of games, for their regurgitation of Hollywood style violence, lack of dramatic credibility or questionable social/political function (McGuire, 2004) (King and Borland, 2003, p.173). The points I have made above about the lack of AI and the unrelenting advance of glossier graphics and spectacular effects would seem to suggest that there is nothing much to be gained from advancing graphics technology capability in games. However, this is not the case.
Brenda Laurel made this clear before the full impact of Moore’s Law saw high resolution gaming hardware in most homes.

First-person sensory qualities are as important as the sense of agency in creating satisfying human-computer experiences. Quite simply, the experience of first person participation tends to be related to the number, variety, and integration of sensory modalities involved in the representation. The underlying principle here is mimetic; that is, a human–computer experience is more nearly “first-person” when the experience it represents unfolds in the appropriate sensory modalities. The intuitive correctness of this notion is witnessed by the directions of technical evolution in the area of simulators and games—towards higher-resolution graphics and faster animations, greater sound capabilities, motion platforms, and mimetic input devices like force-feedback controllers (quoted in Meadows, 2003, p.162).

We are yet to see digitally facilitated games move outside the rectangle of the static screen, and the Nintendo Wii (2006) is the first significant, successful haptic/tactile interface development since the 1970s. However, the improvements in audio and visual rendering have increased “the number, variety, and integration of sensory modalities involved in the representation”. Higher quality images are not necessarily just glossier and more colourful, and richer sound does not just have more bass and come from several different speakers. Richer media do have the potential to actually carry more information, and therefore to convey more meaning for dramatic/narrative/immersive/gameplay effect, even if game designers are not yet fully using this potential.

Further, it would be unfair not to acknowledge that games designers have been using some of the increasing power of hardware for more than just the audio-visual quality. The term “AI” is used very broadly by game designers to refer to improvements in how closely the virtual world can mimic the real world in its behaviour (Rabin, 2004). “Physics engines” give inanimate objects the ability to fall, bounce, break, etc more realistically (Valve, 2004, p.61). Characters can be made to move much more credibly, often with motion data recorded from real life actors (Valve, 2004, p.120). First person shooters’ non-playing characters, the enemy “bots”, are certainly more adept and plausible as combatants since the early 1990s, their design utilising more sophisticated programming (although to keep these developments in perspective, these gradual improvements are probably not yet obvious unless you are very familiar with the first person shooter genre) (Woodcock, 2005).

The capacity for these behavioural advances in the gameworld is, like the audio-visuals improvements, made possible by the advancing processing power of games hardware, and contributes to environments that might move well beyond merely “looking” better.

For example, in the creation of a crime/mystery game, it could be of great significance in the gameplay/narrative if dust falls heavily like sand, if it falls slowly like flakes of dust, or if it floats in the breeze that is stirred by your hand movement. This difference is currently very difficult to recreate credibly. Or it may already be possible, and has just not yet actually been implemented, as designers are distracted with the “sensational”, rather than “subtle” aspects of the audio-visuals.

This kind of subtlety could potentially allow and require a whole different, more contemplative gameplaying aesthetic. Of course, fans of the Myst series might point out that this observational, contemplative gameplay style has already existed. Myst was, in its time, an environmentally immersive breakthrough requiring, and allowing, significant observation skills of the player. Have Myst and its lineage just been too far ahead of their time, despite the early, astonishing success? Will the future history of games design consider that Myst was simply over run by gameplay forms that were less susceptible to a lack of artificial intelligence, faster to incorporate Real Intelligence, and more able to exploit the sheer spectacular potential of early games hardware? Will Myst even be remembered in an ever deepening history of games?

Conclusion
Not only in the Myst and Doom lineage, but across many genres of computer games, we see the inevitable improvements in image quality and real time rendering clearly narrowing the gap between these landmark games’ original aesthetics. But we can also see the parallel lack of improvement in AI, either in characters or storylines (Newman, 2004 p.101). We now have the visual rendering capacity to make the faces, to make the faces animate, and to give the faces beautifully back lit hair, but the faces still cannot have a conversation with you. Computers are much, much faster since 1993, but not much cleverer.
Images and sound of increasing quality, and the associated improvements in physics engines and other behavioural aspects of virtual environments, actually offer up a far greater range of mimetic possibilities than now exist, because they can carry more actual information. This alone opens the possibility for a broadening of game aesthetics, and evolution of genres.

But what happens if gaming *audio-visual* rendering does reach a plateau? Perhaps this would prompt a resurgence in AI communications and interactive storytelling as games designers compete to win the favours of gameplayers. This might fulfil the lost potential that Aarseth noted in 1997 (Aarseth, 1997, p.128). These “clever” development may become the new “spectacular” that fuels a further cycle of player’s expectation and designer’s craft.

There are already some encouraging signs of an “all inclusive” evolution. For example, in 2005, Cyan Worlds released *Myst Five*. The game includes a form of language recognition where players can draw symbols onto a special device, causing the gameworld’s characters to change their behaviour in response (Logan, 2005). While this is still a long way from natural language recognition, and may not even be much more sophisticated than teaching a hand held device to recognise your handwriting, it is an innovative and welcome attempt at “communications AI” in mainstream games. Also of note in 2005 was the release of Andrew Stern’s and Michael Mateas’ interactive drama *Facade*. This strongly research informed experience has re-energised the discussion about natural language recognition, incorporating it into an environment with rich audio and visuals, and additional interactive inputs (Rauch, 2006, p.82).

Whether or not *Myst* and its kind rise once again, games and culture in general will always owe a great debt to *Doom* and its spawn. The development of virtual blood and brains that spatters more realistically has created the “games engines” that are possibly better named “mimesis engines”, especially since they may eventually incorporate communications AI. When put to work for other purposes, these means of mimesis will find all kind of applications in other, future game aesthetics/genres, and also art, education, training, sport, and who knows what else (Woods, 2004).
Non book titles cited


Twentieth Century Fox. (1986). *Aliens*. Feature Film.


Works Cited


http://www.justadventure.com/reviews/MystMaster/MystMaspiece.shtml  


http://gamestudies.org/0601/articles/montfort  


http://money.cnn.com/2004/02/02/commentary/game_over/column_gaming/  


http://Myst5.com/oldnewsuru.html  


