Scenario: Co-Evolution, Shared Autonomy and Mixed Reality

Dennis Del Favero

iCinema Centre for Interactive Cinema Research, The University of New South Wales

Timothy S. Barker

iCinema Centre for Interactive Cinema Research, The University of New South Wales

ABSTRACT

This paper describes the Australian Research Council funded project *Scenario,* conducted at the iCinema Research Centre, University of New South Wales. Realized through interdisciplinary research, involving the domains of machine learning, interactive narrative and new media art, *Scenario* creates a mixed reality (MR) environment, surrounding the user within a 360-degree stereoscopic space, in which she can interact with digital characters that have a level of autonomy. Through its discussion of the aesthetics and technological architecture of *Scenario* the paper enters into an explanation of what is termed 'co-evolutionary' narrative, a function of the interactive relationship formed between a human user and an autonomous digital character. Understanding interaction as co-evolution the *Scenario* project enlists the philosophy of Gilles Deleuze and Manual DeLanda in order to propose interaction as a dynamic two-way process. The paper begins to unpack this theoretical framework for understanding interaction.

Keywords: Mixed Reality; Media Art; Interactivity; Interactive Aesthetics; Machine Learning

Index Terms: J.5 [Computer Applications]: Arts and Humanities --- Fine Arts; H.5.1 [Multimedia information Systems]: Artificial, Augmented and Virtual Realities

1 INTRODUCTION

Digital aesthetics, in broad brush strokes, can be described as the critical reflection of the way digital technologies intersect with, or impact upon, the sensory processes by which we come to know the world. However, this is not to say that digital aesthetics are concerned merely with the 'representation' of the world by digital means. After all, new technologies are used most powerfully when they are applied as constructive, rather than 'representational' or decorative forces, expressing new ways of living in, understanding and interacting with the world. In this sense, we can say that at the heart of digital aesthetics is a concern with understanding the meaningful interactions between technological and human processes. Mixed Reality (MR) environments are progressively upscaling the everyday convergence of the technological and the human, dissolving the

d.delfavero@unsw.edu.au
t.barker@unsw.edu.au
iCinema Centre for Interactive Cinema Research
University of New South Wales
PO Box 259, Paddington, NSW, 2021
Australia

usual boundaries found at the interface of the screen, as digital images can enter the user's space. In these cases, and particularly where digital characters are given the capacity to act autonomously, interaction might be thought to embody an *ecological* relationship between users and digital characters, each affecting the way the other behaves and interprets the world. The relationship is ecological in the sense that interaction is a result of the interdependency between the human and the digital system, re-formatting the conventional aesthetic distinction set up between a beholder and a representational system [1]. Rather than a beholder encountering a linguistic or plastic 'representation', they, as agents in an interactive system, encounter other (digital) agents, such that both share levels of agency. Aesthetic considerations of such a model would focus on the philosophical concept of practical reasoning – as both human users and digital characters consider the way to act, or what to do, based on a relationship formed between internal and external forces – and a theory of narrative as central to the way we make meaning of the episodic events of such a digital encounter.

This paper describes the Australia Research Council (ARC) funded project *Scenario* (2005-10) (fig. 1), undertaken at The University of New South Wales' iCinema Research Centre. The ultimate goal of the *Scenario* project was to test a theory of interaction in which human users and digital characters that are provided with a level of autonomy, are able to share agency in creating an interactive narrative. By this, the research returned an interactive artwork that experimentally demonstrates this co-evolutionary narrative theory, allowing users to enter a MR space that is populated with digital characters and to unfold a narrative of events in co-operation with these characters. This concept will be explained in the paper and offered as a new aesthetic approach to understanding interaction, one in which narrative agency, rather than being solely possessed by the user or the artist, is instead theorized as shared between the human and the machine.

Founded in 2002, iCinema is an interdisciplinary research centre, bringing together researchers in the fields of new media art, computer science and engineering, multimedia design, cognitive science, aesthetics and cinema theory. Its research, realized through artistic, experimental and theoretical pursuits, is primarily focused on the aesthetic use of interactive technologies within immersive cinematic contexts and the way that this can be used to reformulate our day-to-day practices.

Scenario, with its focus on testing the meaningful relationships that may be formed between humans and technology, utilizes innovative research in the field of machine learning and artificial intelligence (AI), along with iCinema's ongoing research into immersive and interactive environments. In this paper we will firstly describe the visualization and interaction infrastructure that makes the *Scenario* research possible and then secondly enter into a specific explanation of the *Scenario* project. The work itself establishes several advances to the field of interactive narrative research within MR environments and to the theoretical frameworks of interaction. Most important is a concept that has been termed *co-evolutionary narrative,* a model of interaction in which a narrative of events is formed based upon the transaction between a human user and a digital character. Theoretically framing this model of interaction with the philosophy of Gilles Deleuze and Manual DeLanda, we understand interaction as an

ecology or an assemblage, involving multiple parts working *with* and working *on* one another, marked by the meeting of human and digital processes within the MR space. After a discussion of the *Scenario* project we will enter into an explanation of this theory and gesture toward the ways that it can be used to understand the relationship between humans and technology, particularly in terms of MR.

2   DISCUSSION

2.1   Beyond the narrative script

Upon entering the dark installation space, filled with the audio provided by a slow moving piano composition, the user, along with four others, sees computer generated eyeballs surrounding her, arranged across the panoramic screen. A voice-over commands the user to step closer to an eye, to choose one; otherwise the eye will choose them. Once this initial act is complete the screen returns to darkness and we hear the sound of rain above us, falling into gutters, and we see a dull half-light, coming through windows overhead. There is a felling that we are in the darkness of an underground maze, at once puzzling due to the ambiguity of the sensory objects that surround us but also in some senses relatively familiar as we can see our physical bodies and the bodies of the other users. A voice from behind, or next to us, depending on our physical position in the space, belonging to a digital humanoid character tells us to come closer, leading us around the panoramic screen that is the boundary of the mixed reality space. They lead us down corridors, showing us body parts; though not from a human body, these are the smooth textured and 'bloodless' body part of a digital character. Here we activate the pre-scripted narrative of a crime based on our movement toward the screen; narrative information is silently unfolded for us as we approach specific characters. All the time, as we move down each corridor, accompanied by the audio of news bulletins and with images appearing to enter our space due to the polarized 3D glasses, shadow characters are ever present, lingering in the dark corridors. After we have navigated these scenes, the narrative space shifts to a clearing in the woods, with snow gently falling. The five digital characters now beckon toward us to help them locate the body parts that are strewn throughout the space. Always attempting to avoid the shadow characters, who attempt to block our actions, we help them reassemble the figure of a chid, who, when complete again, walks through the snow, out of the scene.

The interactive task set out for us in the third act, simple at first, becomes increasingly difficult as the further we move into the space the more the digital characters encounter a cohort of shadow characters who block and undermine their actions. The only way to circumvent these shadow characters is to outsmart them. To do this, we need to find ways of moving to counteract and outflank the shadow characters. This is a complex undertaking as the shadow characters are able to interpret our actions, and intelligently respond to each move that we makes. It becomes apparent that the shadow characters are not simply following scripts, but rather 'deliberating' on their actions based on the situation and how we move in relation to them. Once the child and the characters begin to leave, the piano piece with which the work begun recommences and the space becomes filled with this subdued soundtrack and the snow slowly falling all around us, now appearing much heavier. In this narrative we are not simply beholders in an immersive space, although in the second act this is primarily our role. Instead our role shifts from a passive beholder to an active user, as we attempt to simultaneously *perform* and *understand* the events, becoming conscious of the way our movement is registered by a machine whilst attempting to test out and cognize the consequences of our movements on the aesthetics of the narrative.

Figure 1: Dennis Del Favero, Jeffrey Shaw, Steve Benford, and Johannes Goebel, *Scenario,* 2005-10

In *Scenario* the capacity for digital characters to respond to situations in a largely unpredictable and relatively autonomous manner is achieved by the use of a symbolic language, the operation of which is similar to the theory of affordance developed by the American psychologist J. J. Gibson [2]. In affordance theory, just as in *Scenario,* the world is not merely made up of objects and their spatial relationship, instead the world is composed of the object's potential for action or affect, understood, following DeLanda, as its capacity to act on other objects [3]. For instance, a rabbit chased by a fox views the situation in terms of affordance. The fox affords the rabbit predatory danger, possible death if the rabbit does not flee. The burrow, through the rabbit's eyes, affords a place to escape the pursuit. For the fox, the rabbit affords food and the burrow affords the fox the loss of food.

This is similar to the functioning of machine learning in the third act of *Scenario*. Here physical gestures performed by the human user afford the digital characters various things. For instance if a human user advances towards a digital character, the character may attempt to avoid or hide from the user. The planning system forms a theory of action based on situation calculus. It uses a cognitive robotics programming language based on the Golog family of languages suitably extended to deal with the particular environment imposed by *Scenario*. Roughly speaking this language endows each digital character with its beliefs about the world in which it finds itself and allows it to reason about a proper course of action to take in response to the occurrence of external events.

Scenario functions by sensing the movements of a human user, interpreting this behavior symbolically and enabling digital characters to respond via a system of scripted 'beliefs' and 'desires'. Machine agents respond to actual events in a somewhat unscripted but intelligible way. AI and machine learning research, although still only able to give these digital characters a modest level of autonomy, allows interaction to bypass the borders traditionally experienced between the human and the digital and between the conventional narrative script and autonomy. These developments allow the space to become truly *mixed*: the physical space is not only populated with stereoscopic images, but these images are able to interact with the user within this space, not through scripted encounters, but as both the user and the digitally generated characters take on the role of protagonists within the interactive narrative.

2.2   The mixed reality space of *Scenario*

Scenario takes place within a MR environment made possible by the technical architecture of the Advanced Visualization and Interaction Environment (AVIE) (fig. 2)*,* co-designed by Jeffrey Shaw, Dennis Del Favero, Matt McGinity, Ardrian Hardjono and Volker Kuchelmeister, at the iCinema Research Centre [4]*.* This environment is comprised of a 360-degree projection

environment, a cylindrical silvered screen four meters high and ten meters in diameter, capable of rendering 3D moving images and surround sound. Upon entering AVIE the user, wearing polarized 3D glasses, is surrounded by stereoscopic images that can move freely through the physical space due to a unique configuration of a set of twelve projectors in six stereoscopic pairs, each pair supplying sixty degrees of the whole image. The images that seem to move through the space are also accompanied by audio made possible by a twenty-four channel system with custom surround audio application software. This system enables fully immersive 360-degree placement of sound anywhere around the viewers. As such, the user's aesthetic experience is mutually constituted by sound and vision. AVIE's subsequent multi-modality contributes to the feeling of presence within the immersive environment and produces a sense of space that is at once 'other', due to the logic of virtual images, but also familiar, as the human user can still see her own body, and move through the physical space, interacting relatively naturally with the characters that share this space.
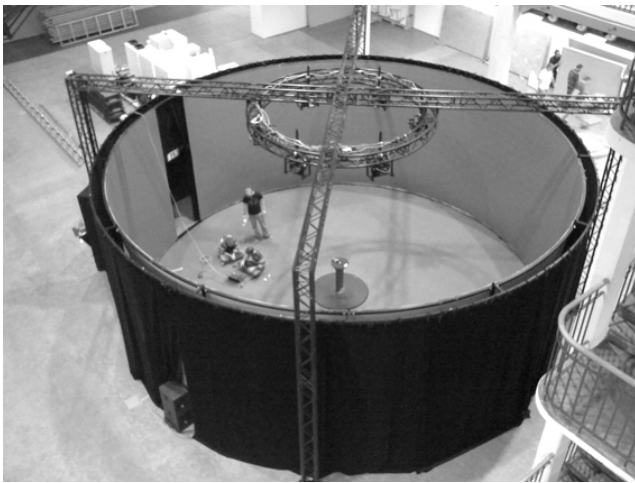

Figure 2: The Advanced Visualization and Interaction Environment (AVIE)

AVIE is also capable of motion and shape tracking, used in *Scenario* to supply the work's necessary gesture detection system. This is made possible by iTRACK, a custom designed system of infrared cameras and real-time software. Both the visual projection and the audio system can be coupled with the Intelligent Vision-Based Interaction and Motion Tracking Systems to allow voices to 'follow' projected characters, or to be activated only when viewers move within 'earshot' of the virtual audio sources.

Figure 3: Dennis Del Favero, Jeffrey Shaw, Neil C.M. Brown, Peter Weibel and Matt McGinity, *T_Visionarium,* 2003-2008

AVIE was developed iteratively throughout previous ARC funded iCinema research. This begun with *T_Visionarium* (2003-2008) (fig. 3)*,* in which over 30 hours of free to air television had been captured and tagged with metadata descriptors to create a database of over 20,000 video clips. In *T_Visionarium* over 300 of these clips are distributed across AVIE's 360-degree screen. Viewers can then reassemble these clips into relational clusters. The clips, appearing in 3D due to the stereoscopic glasses, fly through the space and reassemble themselves upon the surface of AVIE, accompanied by their respective soundtracks. This research saw advances to AVIE's 3D projection system, in particular a method for rendering arbitrary three dimensional scenes in AVIE with globally correct stereo separation, allowing any number of people to perceive correct depth information regardless of their orientation. It also developed an ambisonic decoding engine for playing spatialized and synchronized audio over the twenty-four channel sound system. AVIE has since been utilized as a MR laboratory for the aesthetic research undertaken at the iCinema Centre and is now being brought to maturity through projects such as *Scenario*.

2.3   Performing *Scenario*

The *Scenario* experiment utilizes performance schemas, established by Samuel Beckett in *Quadrat 1* and *2* (1981), an experimental television and theatre work that explores group autonomy based on a series of changing spatial relations. In Beckett's work performers – first one, then building up to four – navigate a rectangular space, pacing around the edges of the space, then moving towards its centre, changing their route to avoid the centre of the rectangle, and indeed each other. The performance is based on a series of planned movements, with performers entering and exiting the space at regular intervals, each one walking to a different rhythm, accompanied by individual percussive soundtracks. The result is a collective of rhythmic pulses, resulting in a singular composition, based upon the group autonomy of the performers.

Beckett's research is employed as it provides an appropriate aesthetic definition of group autonomy as other-intentional, that is, predicated on shared positional actions. In *Quadrat*, for example, characters mutually define each other by means of their respective territorial maneuvers as they move backwards and forwards across the boundaries of a quadrant. *Quadrat* is drawn on as a way of aesthetically conceptualizing the relationship between spatialization and group consciousness. For example, in

one sample experiment for an early iteration of *Scenario*, participants are confronted with several machine agents who are trying to cluster themselves into a group. One of the agents takes on the role of herd master, circling his colleagues as he barks orders. Despairing over his failure, he moves towards the human participants in confusion. The participants in turn may attempt to assist by forming into a group. The agents react unpredictably, possibly mimicking, murmuring and then fleeing the scene. The experiment generates a cascading series of gestural and clustering behaviors, testing and evaluating the network of meaningful deliberations and ascriptions of machine and human agents as they attempt to make sense of each other's behaviors.

The use of gesture recognition in interactive artworks, such as *Scenario*, to place the user's physical body in direct contact with technologically mediated aesthetics, has a rich lineage. Since Myron Krueger's 1969 work *Glowflow*, a light and sound installation that is controlled by a viewer's movements within an installation space, and his more well known 1974 work *Videoplace*, media artists have deployed gesture recognition for artistic purposes, facilitating an embodied, processual and performative sense of aesthetics. This occurs in Krueger's *Videoplace*, for instance, as it captures the motion of participants in two separate rooms via cameras, interprets them as high contrast images and integrates them in a cinematic environment, where the two sets of participants can communicate via their gestures in the shared space of the screen.

We can also see early examples of the artistic use of motion detection in Simon Penny's paradigmatic robotic works *Stupid Robot* (1985) or the later *Petit Mal* (1993). In *Stupid Robot* a robot sits on the gallery floor and detects the motion of visitors as they approach it; when they come close enough, the robot, similar to a beggar, shakes a can of metal parts. In *Petit Mal*, a robot built of a pair of bicycle wheels driven by a motor, coupled with pendulums, body heat detection sensors, processors and a power supply, senses objects and people around it and follows them around its installation space. In these works people interact with machines not by strapping on sensors or manipulating touch screens or keyboards. Instead, people can enact their agency on the machine through unencumbered bodily movements.

Works such as Peter Weibel's *The Wall, the Curtain (Border, the)* (1993) (fig. 4) carry this paradigm forward, using gesture recognition to dissolve the border between the viewer and artwork. In this work a viewer's gestures, captured by a camera, visibly distort images of cave paintings from Lascaux. As the viewer steps in front of the screen to view the projected image, she (after a short time delay), visibly obscures the image, making an uninterrupted viewing of the screen impossible, and thus situating the viewer and the act of viewing as inextricable to that which she observes.
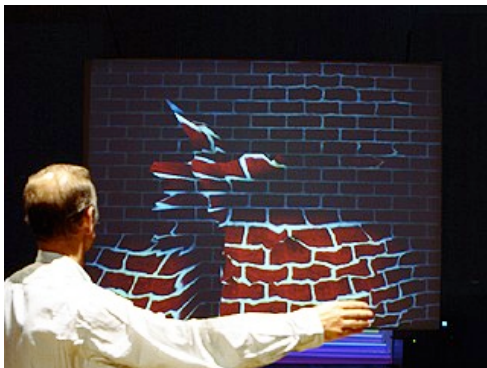

Figure 4: Peter Weibel, *The Wall, the Curtain (Border, the),* 1993

We can also see this occurring in Wolfgang Müench and Kiyoshi Furukawa's *Bubbles* (2000) (fig. 5). In this work a participant can interact with digitally generated bubbles via her shadow cast onto a screen. By stepping in front of the projector's light beam a participant can cause the bubbles that slowly float down the screen to bounce off her silhouette, behaving as if her shadow were a solid object. In this work the body, sensed digitally, becomes central to the interactive event. More precisely, the outline of the body becomes central as an interface to the two-dimensional space of the projection screen. Another example of this is Christa Sommerer and Laurent Mignonneau's *IntroAct* (1995) (fig. 6), a work in which the user's gestures cause the growth of digitally generated organic-like material. As the viewer enters the installation her image is projected onto a large screen, as she moves in certain ways she causes the cinematic space of the screen to become populated by more and more digital forms. For instance, moving her hand a certain way may cause the growth of digitally generated forms out of her palm. In these types of works the feeling of a shared space with the digital is enhanced through the concept of embodiment. The user does not interact by making choices via the conventional and restrictive input devices of the mouse, keyboard or touch-screen. Rather the user must physical *act*, moving as if in the 'real' world. In this sense the relationship between the digital and the physical is *expressed* or *embodied*, in the Deleuzian sense, via these actions.
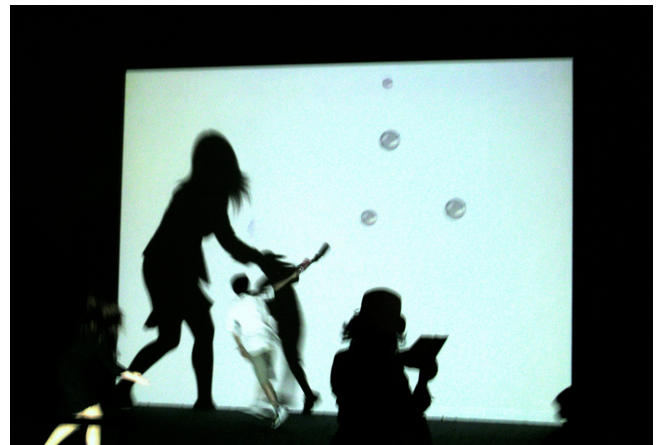

Figure 5: Wolfgang Müench and Kiyoshi Furukawa, *Bubbles,* 2000

Figure 6: Christa Sommerer and Laurent Mignonneau's, *IntroAct,* 1995

Building upon this history, the MR environment of *Scenario* uses a technical architecture that utilizes motion detection as an interface to the digital. However, diverging from the afore mentioned examples, interaction with *Scenario* is more attuned to a real world dialogic interaction, with gestures prompting responses from the digital characters, which then prompt further gestures, rather than triggering or effecting aesthetic images.

The framework of *Scenario* is structured so as to provide a two-way interactive conversation between a real and a digital character in the one immersive MR space. It focuses upon the way an interactive narrative can evolve between a user and a digitally

generated character within this shared space, as opposed to the simulated intentionality of conventional interaction, where a user and the digital can only converse along pre-scripted pathways, and where interaction is focused upon the 'user'. Instead, the type of interaction produced by this research project focuses upon the actual experience, relationships and processes of interaction, as an event involving both human and machine generated protagonists. Instead of focusing exclusively on human judgments, the narrative of events that is interactively produced via *Scenario* includes the contributions made by fictional agents to unfolding events. This contrasts with traditional narrative agency, for example as found in cinema. Traditional agency is planned into the script and post-production process and its impact can be thoroughly brought to light only through the retrospective application of critical interpretation [5]. The impact of agency in *Scenario* by contrast involves a more practical reasoning enacted on the fly, as users and digital characters decide on a plan of action and develop a theory of the world that emerges from the pattern of interactive events.

In terms of a historical lineage of the aesthetic use of AI and machine learning research, we can point to several important instances throughout the history of media art. For instance, we could also term a work such as Ken Feingold's *Head* (1999) (fig. 7) as a co-evolutionary narrative, though not in the MR sense. *Head* is a realistic model of a disembodied head able to recognize and respond to human speech. In this work the participant is able to ask the robotic head questions and carry on a conversation, with varying levels of coherence. The computer runs several software applications, including commercial speech recognition software and an application that attempts to evaluate the meaning of the processed speech and formulate a reply from its database of words [6]. This database is built up by inputting texts into the head's database, in a sense building the machine's vocabulary of known words and sentences. The work here, using these natural language processing routines that Feingold calls an Absurd Conversation Engine (ACE), is capable of conversing with the user, and further producing rhyming responses and alliteration based on what it has 'heard' [7]. In this conversation the user does not completely direct the machine, rather the machine transacts with the user in unforeseen ways, both machine and user enacting agency over the interactive narrative, as the machine attempts to parse the user's speech and as the user tries to make sense of the machine's speech. Interaction here constructs a semi-lucid, but somewhat 'schizophrenic', conversation. The head responds to the questions asked of it. But these responses make little sense, constrained by its vocabulary and its capacity for understanding what is asked of it. For instance, a typical transaction between a visitor and the head may be:

Visitor: How are you?
Head: I am so offended. I can't tell if everyone would prod her to the left or not.
V: Who offended you?
H: The delicious conditions delay disturbed daredevils.
V: Which daredevils are disturbed?
H: The bothersome one
V: Which one is bothersome?
H: The absurd one [8]

Figure 7: Ken Feingold, *Head,* 1999

Although this exchange is somewhat intelligible, in the sense that the user can see that the head is replying in response to its understanding of her questions, the response is nonetheless relatively nonsensical in the context of a conversation. The user responds to these coherent yet distinctly 'other' utterances by asking more questions. In this sense, the head is not 'used' by a user, as the sole possessor of agency directing it to pre-determined courses of action, rather it responds to the speech initiated by a user as it autonomously enacts its own agency.

We can also see the multiple levels of agency at play in Alexa Wright and Alf Linney's more recent aesthetic use of AI research in their work *Conversation Piece* (2008) (fig. 8). In this work exhibition goers encounter two small flesh colored sculptures. As the visitors enter the space they are automatically tracked using webcams. As they pass in proximity of the sculptures the voice of the room attempts to engage them in conversation. Due to the positioning of focused directional speakers, the sound is directed so that only viewers standing near the sculptures can hear the voice. If the viewer replies to the voice's initial 'hello' or 'excuse me', the voice, referring to herself as 'Heather', uses keywords to formulate a reply. Along these lines narrative agency is shared as if, similar to *Head's* back and forth, a real world conversation, with both parties capable of acting on one another and sharing autonomy. Of course here the machine, as with Feingold's *Head,* does not comprehend the meaning or affect of its reply, as we would normally assume in traditional models of intelligence. But this is not the point; in theorizing this type of interaction in which autonomy is shared, as with *Head* and *Scenario,* we consider the machine as an agent, focusing on its capacity to affect and be affected, on the same level as the human user [9].

As the artists point out *Conversation Piece's* interactive environment, similarly to Feingold's *Head,* has its beginnings with Joseph Weizenbaum's *Eliza Chat Bot*, a computer program designed to emulate a Rogerian psychotherapist. Designed in 1966, this work is an early example of natural language processing routines. Eliza operates by parsing inputted speech and substituting key words into pre-scripted phrases. As such, Eliza mostly either rephrases the user's responses and poses them as questions, or enacts strategies to prolong the conversation (which seems rather effective as a system masquerading as a Rogerian psychotherapist). The important point in the research surrounding *Eliza Chat Bot*, *Head, Conversation Piece* and *Scenario* is that it is not necessarily important that the machine actually possess high level intelligence for interaction to take the form of shared autonomy; after all, in order to possess agency one does not necessarily require the capacity to reflect on the use of this agency. In all these works, as pointed out above, the machines do not possess the capacity to reflect on the various meanings of their responses, they merely register the observable consequences of these as a part of the changing world in which they find themselves. Nevertheless, this does not constrain the machine's agency, as its capacity to affect and to act in the world and to enter into a conversational narrative with a user.
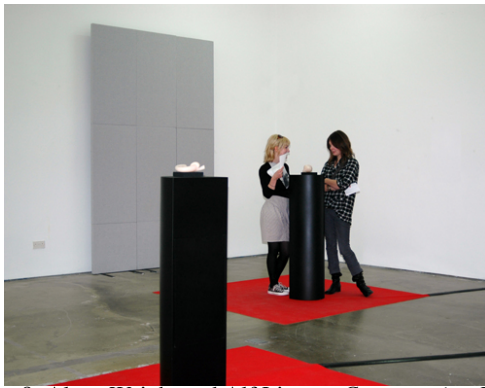
Figure 8: Alexa Wright and Alf Linney, *Conversation Piece,* 2008

As a consequence of this theoretical and aesthetic positioning of intelligence and interactive agency the concept of narrative as an eventful episodic sequence of events is central to our theory of interaction and the *Scenario* experiment. As David Herman has argued in terms of human society, the affordance that the subject infers from situations is based upon their insertion of the events into a narrative structure and the prediction of the outcome, similar to the fox and the rabbit example provided earlier by Gibson [10]. Herman gives the example of an individual advancing toward him with fists raised as data that affords him danger [11]. It does so because Herman has inserted the individual's overt actions into a narrative sequence by which he is able to predict the likely outcome should he not take action. As such, interaction takes the form in *Scenario* as a multi-perspective narrative, with both the user and the software converting the events into a narrative of interaction, attributing affordance to particular actions and objects, and acting based upon their narrative predictions.

2.4 Co-Evolutionary Narrative

A *co-evolutionary* narrative is so termed because the narrative evolves or emerges based on a relationship formed between the actions of a human user and a digital agent. This is a system in which a digital agent is able to respond autonomously to a human user, and it is through this response that a narrative is performed, with both the user and the character sharing narrative agency, expressed as their individual capacity to act on one another in the interactive encounter. User initiated processes and machine initiated processes – as they sense, parse and respond to the user initiated processes – construct the narrative on the fly. They evolve the narrative through a *common operation* within the MR space.

It is as Andrew Stern pointed out in 2001, "it is when artists combine the computer's capabilities of real-time autonomy and reactivity that they achieve a deeper form of interactive art. By making the computer *listen* to the audience (the first half of reactivity), *think* about what it heard (autonomy), and then *speak* its thoughts back to the audience (the second half of reactivity), the artwork can have a dialog, a *conversation*, with the audience (Stern's emphasis)" [12]. By the terms 'speak' and 'conversation' Stern, as we do, does not imply literal speech, but rather any form of meaningful conversation or communicative 'back and forth'. This type of conversational co-evolution of narrative in artworks is rich in aesthetic possibilities due to the work's capability to detect and respond to a particular user, to customize itself and to adapt to the environment in which it finds itself. In this respect the episodic events of interaction that take place in *Scenario's* MR space form a narrative of events. And it is here that narrative is valued as a site for social discourse, for the potential that it has for dialogue across physical and digital spaces, rather than for its internal properties [13].

The importance of the concept of co-evolutionary narrative when applied to interaction, particularly concerning MR research, is that it respects the levels of agency of both the human user and

the digital machine, viewing the aesthetic judgments made in the digital encounter as directed by the transductive relationship formed between these two spaces. For MR environments to bring together real and digital spaces such that interaction can occur between the spaces, one must acknowledge the creative role of the digital, as it attaches itself to our sensory experiences. If a theory of interaction is to be focused solely on the experience and agency of the human user, our understanding of the aesthetic impact and process of MR is severely undercut. For a space to be truly 'mixed' between the digital and the physical, a theory of interaction must be equally focused on both the digital and the physical, allowing the digital to affect the processes of the physical, through interactive deliberations, just as much as the physical can affect the processes of the digital.

Adrian Mackenzie has described this previously in terms of transduction. Using Gilbert Simondon's example of the growth of a seed crystal suspended in liquid, Mackenzie shows how technology both restricts and enables our behaviors as it intersects with culture with increasing ubiquity. The planes on which the crystal grows are always on the crystal's surfaces, in contact with the liquid [14]. Because of this, the process of individuation, or differentiation, is a process that is not initiated by the crystal alone, but rather a process of transaction with its milieu. In other words, the crystal's growth, or its becoming, takes the form that it does due to its own internal energies and potentials put into contact with the external forces of the liquid. The shape of the crystal is thus produced as a commingling of forces of potential and environmental conditionings. Simondon applies transduction here to physical growth, but we can also apply it to other aspects of life, as indeed Mackenzie does, in order to understand it as a dynamic negotiation between forces. In this case we can apply it to the negotiation between human and technological forces, understanding the interaction of human and technology as a becoming that is brought about by a differentiating exchange of energies. The user, similar the seed crystal, does not know the shape or direction that their interaction within the digital encounter will take. It must rather be actualised by a process of negotiations between her own capacities, the interface's parameters, the software's programming and the environmental circumstances encountered in her interaction with the digital. This is all fairly obvious in traditional versions of VR, such as those models that require the full head mounted display, along with data gloves or wands and sensors strapped to the body to track motion. These devices, whilst enabling interaction within the digitally generated environment, significantly constrain the types of movements able to be made in physical space.

Mackenzie shows how the technicity of an ensemble is always *in situ*, it is always localized and encumbered; as such it is not only the wired user that is weighed down by technology. For instance, he points out that the mobile phone or wireless technology is in fact massively constrained by its reliance on an ensemble of networks and communications infrastructure [15]. The point is that technical mediation is always connected, and sometimes weighed down, by its context; it is contingent upon the other agents that constitute its ensemble. But this need not be a negative force, as Mackenzie points out, it is enough to simply say that this is a differentiating force that may both restrict *and* enable. In terms of an interaction with technology, "technologies are not a domain exterior to human bodies, but are constitutively involved in the 'bodying-forth' of limits and differences" [16]. Understood transductively, interaction is a process in which the human is supplemented by the digital in a common operation that conditions the becomings of both entities.

For instance, in *Scenario* the movement of the human user, performed in physical space, are sensed via iTRACK and communicated to the digital characters, which then reason about an appropriate course of action to take. Seen here, interaction occurs as a transduction, as the user's movements are restricted by the technological infrastructure of *Scenario,* not in the sense of strapping on sensors or wearing full head mounted VR displays, but rather in the more subtle sense of requiring that the users regulate their movements to the 'rhythm' of the machine and to the 'rhythm' of the narrative. For instance, we have observed that users tend to move in *Scenario* in a much slower and deliberate

manner than in real world interactions. This may be due to a number of things, but we may frame this, following Mackenzie, firstly as a consequence of the immersion in such a subdued and foreign narrative space, established by the audio of the piano piece and the darkness of the space, from which the humanoid characters emerge. Secondly, the users' movements are affected as they attempt to regulate physical movements to the movements of the characters on the screen, as they follow the users around the space. Thirdly, because the users are innately aware that they are being closely watched and that all of their movements are being given significance, they may tend to reason more thoroughly about the consequences of their otherwise 'natural' movements, which produces these slow, deliberate movements, largely designed to 'test' their effect on the digital characters.

Figure 9: Dennis Del Favero, Jeffrey Shaw, Steve Benford, and Johannes Goebel, *Scenario* (showing curve of AVIE), 2005-10

2.5 DeLanda, Deleuze and Assemblages

The model of interaction for *Scenario,* as well as the other examples mentioned throughout this paper, can be framed by the assemblage theory developed by Deleuze, and expanded by Manuel DeLanda in his book *A New Philosophy of Society*. For DeLanda assemblages are a way to discuss a collection of wholes, such as humans, societies, organs, atoms, or ecosystems, and, for our purposes, can be used to understand interaction. Here an assemblage is made up of parts, however, it is always more than a mere aggregate of these parts. For instance, it is not that a human body is constituted simply by an aggregate of organs. Rather, the human body is constituted by the organs' capacity to act and to work with one another. Similarly a society is not made up merely by a sum of people. What makes the society an assemblage is the capacity that each individual has to interrelate within the collective. In short, an assemblage is always constituted by the capacity for interaction between its parts. We should thus think of an assemblage as constituted by the capacity, or the potential for action [17]. It is the *affect* of the parts – as their capacity to act on one another – that matters, not their materiality, individual power or visual appearance [18]. For instance, we could think of the internal combustion engine as an assemblage. All of its parts have a set function within the assemblage, and the running of the engine depends on each of these parts carrying out its function. If a rubber hose fails to carry the fuel to the engine then the car lies useless on the side of the road, irrespective of the potential for power immanent to the engine [19]. The assemblage here is not constituted by the physical qualities of the carburetor, the cylinders or the crankshaft; it is rather constituted by the potential for the carburetor, the cylinders and the crankshaft to work together by carrying out their individual routines. In other words, the assemblage is constituted by the way its parts *act* within the assemblage, not necessarily by any given fact or trait of the parts [20]. In this sense we can understand interaction as an assemblage. Importantly in *Scenario* this is an assemblage across the MR space of the real and the digital. In *Scenario,* what we can now term the 'interactive assemblage' (which is somewhat of a tautology, as all assemblages are interactive), is constituted by the potential for action expressed by the human user and the digital

characters, as each move in a way that is governed by the assemblage that they find themselves within.

DeLanda gives real world examples of assemblages in action that are prescient to the model of interaction in *Scenario*. Firstly he gives the example of a human conversation as an assemblage, pointing out that human conversation is made up of a set of rules and conditions which are constituted by the flow of words and specific organizing principles [21]. The assemblage here thus involves people, language and many other acts that support, develop and stabilize the friendship (acts such as having dinner together or watching one another's children).

DeLanda's assemblage is embodied in *Scenario* as an assemblage is formed between a human user, a group of modestly autonomous digital characters, the software routines that actuate these characters, the hardware of AVIE, computer programmers and the work's authors. These parts of the assemblage each have specific roles that they play out in order for the assemblage to function. For instance the human user must move through the environment, interacting with the assemblage as they are sensed via the motion detection system, the virtual characters then must be able to sense this action, interpret it as an external event and take a course of action depending on the set of capacities embedded in their programming. Here there is a set of rules that the assemblage functions within, and a technical and aesthetic architecture that stabilizes the assemblage and allows for meaningful interactive relationships to be formed. For the interaction between the user and the digital characters to function the hardware of AVIE must play out its interactive function, the motion detection system iTRACK, for instance, must detect movement and translate this into digital information, enabling the parts of the assemblage to interact and to actuate their capacity to act on one another, providing an interface between the human/digital boundaries, just as the rubber hose carries petrol to the engine.

The assemblage is thus constituted by the conditions imposed by both the internal programming and limitations of the computer as well as the compositional processes initiated by an artist, the technical processes initiated by a computer programmer and the interactive processes initiated by a human user. In other words, it is the process of the artist, the computer programmer, the software and the user that provide the condition from which an interactive narrative emerges [22]. The output of interaction, whether this is an image on the interface, an affective impulse on the user, or a physical movement, thus *expresses* or *embodies* the attributes of the assemblage. To understand how any collective, assemblage, or machine, including the digital machine, is able to produce new or novel information we need to understand the potential for action embedded in every element of the assemblage as a field of emergence, a field or grounding that conditions the manner in which novelties actualize [23]. Our use of Deleuze and DeLanda to understand the deployment of AI within the MR environment of *Scenario* may enable us to understand MR, machine learning and AI aesthetically and philosophically, outside of their usual technical definition.

In *Scenario,* as already discussed, there is an acknowledgement of the various levels of agency involved in the construction of interactive narrative. This involves the agency of an artist, the agency of the computer programmer, the agency of the software and the agency of the user. In this way, the framework put forward here is focused on a techno-aesthetic theory, it is focused on providing an understanding of digital aesthetics and interactive narrative not merely based on the appearance of the screen interface or the experience of the user, but rather based on the process of the system, and its potential for action, including the usually invisible layers of software and the interactive processes and relationships that perform the interactive artwork.

3   CONCLUSION

Based on the use of AVIE, the deployment of AI and machine learning research and the emphasis on interactive aesthetics as a performative and co-evolutionary interactive narrative, *Scenario* crosses boundaries between the digital and the physical. Firstly, within AVIE, the boundary of the screen collapses as the human user is surrounded on all sides by stereoscopic images that seem to populate the same physical space as her own body. Secondly, by using gesture recognition, without the need for restrictive or cumbersome interactive devices the boundaries traditionally set up by an interface are disintegrated. Conventionally in interactive platforms there is a distinction between the user at the controls and the digital characters responding to these directions. Instead here a narrative is co-evolved as both the user and the digitally generated characters become protagonists in the interactive narrative that autonomously takes shape in the digital encounter within AVIE. Thirdly, through the use of AI and machine learning the conventional boundaries between the user, as the sole arbiter of autonomy, and the machine are traversed as digital characters now share autonomy within the narrative. Here, it is not so much a 'user' that uses a machine, but rather a transaction between two entities within one MR space. This borderless interaction, where agency is shared between the entities of the user and the machine, can be understood as an assemblage or ecology, where the interactive narrative co-evolves as the user and the machine *express*, *perform* or *embody* their relationship.

REFERENCES

[1]   D. N. Rodowick, *Reading the Figural, or a New Philosophy for New Media* (Durham: Duke University Press, 2001).

[2]   J. J. Gibson, "The Theory of Affordances" in *Perceiving, Acting and Knowing,* eds. R. Shaw & J. Bransford (Hillsdale: Erlbaum, 1977); J. J. Gibson, *The Ecological Approach to Visual Perception* (Boston: Houghton Mifflin, 1979).

[3]   M. DeLanda, *Intensive Science and Virtual Philosophy* (New York: Continuum, 2002), 72-73

[4]   M. McGinity, J. Shaw, D. Del Favero, and V. Kuchelmeister, "*AVIE*: A Versatile Multi-User Stereo 360-Degree Interactive VR Theatre" in *The 34th International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2007*, San Diego, USA, 5-9 August 2007.

[5]   N. C. M. Brown, T. Barker, and D. Del Favero, "Performing Digital Aesthetics: The Framework for a Theory of the Formation of Interactive Narratives" *Leonardo* (forthcoming), 2011.

[6]   C. Erkut, "Abstraction Mechanisms in Computer Art." *Seminar on Content Creation*, Helsinki University of Technology, 2000.

[7]   K. Cleland, "Talk to Me: Getting Personal with Interactive Art" in *Interaction: Systems, Practice and Theory*. The University of Technology, Sydney, 16-19 November, 2004.

[8]   From http://www.kenfeingold.com/artworks90s.html

[9]   B. Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory* (Oxford: Oxford University Press, 2005).

[10]   D. Herman, "Narratology as a Cognitive Science," *Image and Narrative*, 1 (2000 [cited 8 April 2010), http://www.imageandnarrative.be/inarchive/narratology/davidherman.htm.

[11]   Ibid.

[12]   A. Stern, "Deeper Conversations with Interactive Art, or Why Artists Must Program," *Convergence: The Journal of Research into New Media Technologies* 7(1) (2001):17-24

[13]   S. Cubitt, "Spreadsheets, Sitemaps and Search Engines: Why Narrative is Marginal to Multimedia and Networked Communication, and Why Marginality is More Vital then Universality," in *New Screen Media: Cinema/Art/Narrative*, eds. Martin Rieser and Andrea Zapp (London: BFI, 2002), 1-13.

[14]   A. Mackenzie, *Transductions: Bodies and Machines at Speed* (London and New York: Continuum, 2002), 17.

[15]   Ibid., 12

[16]   Ibid., 52

[17]   G. Deleuze, *Expression in Philosophy: Spinoza* (New York: Zone Books, 1990).

[18]   G. Deleuze and F. Guattari. *What is Philosophy* (London: Verso, 1991/1994), 66-68.

[19]   J. Williams, "Correspondence Why Deleuze Doesn't Blow the Actual on Virtual Priority. A Rejoinder to Jack Reynolds," *Deleuze Studies* 2 (2008): 97-100

[20]   G. Deleuze and F. Guattari. *Capitalism and Schizophrenia* (New York: Continuum, 1980/2004), 285.

[21]   M. DeLanda, *New Philosophy of Society* (London: Continuum, 2006), 87

[22]   C. V. Boundas, "Ontology" in *The Deleuze Dictionary,* ed. Adrian Parr (Edinburgh: Edinburgh University Press, 2005), 192

[23]   B. Massumi, "Sensing the Virtual, Building the Insensible," *Architectural Design* 68(5/6) (1998).