PROCEEDINGS S I G N 2 0 1

1st Conference on Computational Semiotics for Games and New Media

10 – 12 September 2001 CWI - Amsterdam

ISBN 90 6196 5047 NUGI-code 811

Copyright © 2001 Stichting Centrum voor Wiskunde en Informatica

Printed in the Netherlands

Local Chair	Frank Nack, CWI, The Netherlands	
Organizing Committee	ee Andy Clarke, Kinonet Consultancy, UK	
	Clive Fencott, University of Teeside, UK	
	Craig Lindley, Interactive Institute, Sweden	
	Grethe Mitchell, University of East London and Kinonet Consultancy, UK	
	Frank Nack, CWI, The Netherlands	
Proceedings editor	Frank Nack, CWI, The Netherlands	
Programme Committee	Elisabeth André, Universität Augsburg, Germany	
	JP Balpe, l'Université Paris 8, France	
	Paul Brna, Leeds University, UK	
	Kevin Brooks, Motorola Human Interface Labs, USA	
	Andrew Clarke, Kinonet Consultancy, UK	
	Clive Fencott, University of Teeside, UK	
	Monika Fleishman, GMD, Germany	
	Ian Green, University of Westminster, UK	
	Werner Kriechbaum, IBM, Germany	
	Craig Lindley, Interactive Institute, Sweden	
	Grethe Mitchell, University of East London and Kinonet Consultancy, UK	
	Frank Nack, CWI, The Netherlands	
	Paolo Petta, ÖFAI, Austria	
	Doug Rosenberg, University of Wisconsin-Madison, USA	
	Adelheit Stein, GMD, Germany	
	Ola Stockfelt, Göteborgs Universitet / Högskolan Skövde, Sweden	
New Media		
Artworks Committee	Andrew Clarke, Kinonet Consultancy, UK	
	Maja Kuzmanovic, Starlab, Belgium	
	Mari Mahr, Photographer/Artist, UK	
	Grethe Mitchell, University of East London and Kinonet Consultancy, UK	
	Doug Rosenberg , University of Wisconsin-Madison, USA	
	Marah Rosenberg, Avaya Labs, USA	
	Lillian Schwartz, Lucent Bell Labs, USA	
	Uli Spierling, ZGDV, Germany	
	Linda Wallace, Machine Hunger/Australian National University, Australia	

Table of Contents

Preface	VI
Keynote speakers	VIII
Papers	
New media semiotics – computation and aesthetic function	3
Craig Lindley, Frank Nack, Andy Clarke, Grethe Mitchell,	
Clive Fencott	
Issues of Sequence in Converging Media: Studying World Wide Web Documentaries	9
Anders Fagerjord	
Montage and Digital Semiosis	19
Ramesh Srinivasan, Roderick Ferguson, Befekadu Ayenew	
nTRACKER	24
Margarete Jahrmann & Max Moswitzer	
Cypher: Cyber Photographer in Wonder Space	27
Takahiro Nakai, Makoto Tadenuma, Shoji Tanaka, Keiko Nakao	
The Link is the Data - On Realisations Representations, and > the Link Betwixt	31
Werner Kriechbaum, Gerhard Stenzel	
The Data][h!][bleeding T.ex][e][ts_	41
Mez Breeze	
Seine hohle Form Artistic Collaboration in an Interactive Dance	
and Music Performance Environment	43
Butch Rovan, Robert Wechsler, Frieder Weiss	
Streams of Motion (Super Spectacular) – A Virtual Reality Art Work	48
Timothy Portlock	
Literary Theory and Computer Games	51
Julian Kücklich	
Designing Interaction Narratives	59
Mikael B. Skov and Peter Bøgh Andersen	

Polygon Destinies: The Production of Place in the Digital Role-Playing Game				
Julian Holland Oliver				
Beyond Manzanar – Constructing Meaning in Interactive Virtual Reality				
Tamiko Thiel				
Semiotic and nonsemiotic MUD Performance				
Ragnhild Tronstad				
From "First-Person Shooter" to Multi-user Knowledge Spaces				
Mathias Fuchs, Sylvia Eckermann				
Alternumerics				
Paul Chan				
Bridging the Semantic Gap in Content Management Systems: Computational Media Aesthetics	94			
Chitra Dorai, Svetha Venkatesh				
Introducing Interrelations Of Madness & Virtuality	100			
Steffen P. Walz, Isabel Zundel				
A Missing Link: The Role of Semiotics in Multiagent Environments	108			
Mirko Petric, Inga Tomic-Koludrovic, Ivica Mitrovic				
POSTER				
Interpoetry: a game of words, images and sounds as a poetic sign in				
digital media	114			
Jorge Luiz Antonio, Philadelpho Menezes, Wilton Azevedo				
Narratives Argument for Interaction Models: Or how our unstable				
double-agent found its way into the textual machine	117			
Jaap Drooglever				
Sentient VR Environment Design: The Memesis Project	121			
Alison McMahan				
Metaphors and Multimodal Interaction	125			
Dorothy Rachovides, Zoe Swiderski, Alan P. Parkes				
A semiotic communication model for interface design	129			
Rafaella Scalisi				
Life/Style OnLine: A Web-Based Methodology for Visually-Oriented Consumer Research	132			
Michael R. Solomon, Basil G. Englis				

Preface

The idea for this conference came up two years ago, as a result of long ongoing discussions concerned with the ways in which new media systems encode and convey meaning to system users. It soon became apparent to us that there is a great need for common languages between computer scientists and software engineers, media producers, artists, theorists, and critics. This need is driven in particular by the fact that the communicative, aesthetic, and expressive functions of new media are determined by both the computational code and more obvious audiovisual content of new media productions. New media systems support new forms of expression, and represent a rapidly evolving field for the production of meaning. Semiotics as the science of signification, of the production and intercommunication of meaning, therefore emerges as a strong candidate for the development of rigorous analyses and principles for understanding and creating these new forms.

Unfortunately, it turned out that the major conferences in the field of multimedia and hypermedia were concerned with other issues. We therefore started in the year 2000 with a workshop on computational semiotics for new media, being held at the University of Surrey. The questions we wanted to discuss covered quite a wide range of topics, such as

- applicability and application of semiotic theories to new media
- system analysis and design principles, tools and techniques, software architectures and technologies based upon generic semiotics or models
- models of genre, style and form (eg. film genres such as film noir, romantic comedy, western, etc., computer game genres such as adventure, simulation, RPG, etc.)
- principles for multi-modal design and interaction, cross-modal relationships and metaphors narratology, including both linear and non-linear narrative form within new media
- theories of agency in new media
- role of AI representation formalisms and techniques in new media, such as scripts, conceptual graphs, learning algorithms and evolutionary computing

Despite very short notice a relatively big crowd gathered and the discussion became lively. A particular outcome of this workshop was the understanding that neither a mere computational nor artistic view on the matter would help to solve the challenging problems of a communal culture being hooked into an ever growing media-based information space. However, at the end of the day we had produced more questions than answers and thus we decided to offer a discussion space for topics related to computational semiotics for games and new media. COSIGN was born.

For this first conference we decided to mix academic papers and artistic presentations – yet still allocating enough time for vital and vibrant discussions. We hope that at the end of the three days we know better where to go from here and perhaps depart with the anticipation for COSIGN 2002.

Producing a conference requires the effort of many individuals. We would like to thank all the people who volunteered their time and talents in support of COSIGN 2001; the members of the programme committee and the new media artworks committee, the authors of the papers gathered in these proceedings, and those at CWI who helped us in making Amsterdam the right location for the first COSIGN conference. We are in particular most grateful to Marja Hegt, who heartily managed the whole administration. Moreover, many thanks to Miente Bakker,, who was responsible for the CWI registration environment. We wish to express our gratitude to Niels.Nes, Stephan Manegold, and everybody at CHIP for providing the technical support. Extra thanks to Michèle Huijberts and Inge Brouwer, who helped supplying Palidrome with the appropriate dancing environment for their performance during the conference. Many thanks

also to Jos van der Werf and Jan Schipper for solving all problems with the printing of the proceedings.

A conference also requires an operating budget, which in case of COSIGN 2001 is funded mostly by registration fees. However, through the generosity of CWI the conference was able to moderate registration fees while providing the appropriate environment and organisation for this conference. We gratefully acknowledge this support. We also would like to thank DWA (DansWerkplaats Amsterdam) for supplying us with the dance floor and the lighting equipment.

Finally, we would like to thank the participants of this conference because without you no stimulating interaction would be possible.

Andy Clarke, Clive Fencott, Craig Lindley, Grethe Mitchell, and Frank Nack

Keynote Speakers





Arthur Elsenaar & Remko Scha

BIO

Arthur Elsenaar is an artist and an electrical engineer. He used to run his own pirate radio station, and he built the transmitters for many illegal radio and television stations throughout the Netherlands. His radar-controlled interactive sculptures were shown in several international exhibitions. Elsenaar's recent artistic work employs the human face as a computer-controlled display device. Elsenaar coordinates the New Media curriculum at the Frank Mohr Institute in Groningen.

Remko Scha is an artist and a computer scientist. He has built an automatic electric guitar band ("The Machines"), designed an image generation algorithm ("Artificial"), and developed a language-processing theory ("Data-Oriented Parsing"). He coordinates the Computational Linguistics curriculum at the University of Amsterdam, and performs as a DJ on the Amsterdam pirate station "Radio 100".

Arthur Elsenaar and Remko Scha have jointly developed a series of automatic performance-pieces, video-installations and audio-installations, involving computer-controlled facial expression, algorithmic music, and synthetic speech. These works have been presented at scientific conferences, theatre festivals and art exhibitions throughout Europe and the United States. Elsenaar and Scha are also exploring the use of automatic radio stations as a medium for computer art.

CONTACThttp://www2.netcetera.nl/%7Eiaaa/artex/main.htmlTitleThe Semiotics of Human Facial Expression - a Computational Approach.AbstractIn this talk, we investigate the human face as a sign system. To provide an
experimental basis for this investigation, we have designed a MIDI-based
face-interface. This interface allows a digital computer to trigger a large
variety of muscle configurations on the face of a live human person. We will
demonstrate this computational set-up, and show that a computer-controlled

demonstrate this computational set-up, and show that a computer-controlled human face is capable of many interesting patterns, which are normally not employed by the human brain. We will conclude by discussing various possible applications of this technology.



Ken Perlin

BIO	Ken Perlin is an Full Professor in the Department of Computer Science and the director of the Media Research Laboratory at the Courant Institute of Mathematical Sciences of New York University. He is also the director of the NYU Center of Advanced Technology, sponsored by the New York State Science and Technology Foundation.	
	He completed his Ph.D. in 1986 from the New York University Department of Computer Science. His dissertation received the Janet Fabri award for outstanding Doctoral Dissertation. His research interests include graphics, animation, and multimedia	
	Dr. Perlin was Head of Software Development at R/GREENBERG Associates in New York, NY from 1984 through 1987. Prior to that, from 1979 to 1984, he was the System Architect for computer generated animation at Mathematical Applications Group, Inc., Elmsford, NY.	
	TRON was the first movie for which his name got onto the credits. He has served on the Board of Directors of the New York chapter of ACM/SIGGRAPH, has been a member of ACM and ACM SIGGRAPH, and has been a senior reviewer for a number of technical conferences.	
CONTACT	http://mrl.nyu.edu/~perlin/	
Title	Star Wars Chess on the Holodeck	
Abstract	In Star Wars we all saw Animated Chess pieces cavorting on a chessboard, and Princess Leia projected in nothing but a beam of light. And on some level we realized what a profound impact such technologies would have. But are they possible? If so, how and when? And what about building that Star Trek Holodeck?	
	A good start from the display side is to display 3D stereo images (in which each of your two eyes can see a different viewpoint), without requiring the use of special stereo glasses. Ideally, an observer of such an image should be able to move around and change position freely. In our lab we have built a novel display device that does precisely this.	
	I will also show some existing prototypes for the Star Wars chessmen (and for their near descendents – the chessmen of "Wizard Chess" in the Harry Potter books).	
	But what can we expect in five years? I will outline a research vision to get us closer to the Holodeck -in which different people in the same room can see different 3D scenes upon the same walls, without needing to wear 3D glasses.	
	I will also try to look forward several decades into the future, to ask the question: "What would it really take to simulate Princess Leia in that beam of light?" I claim that not only is it possible, but that we can and should start various research steps right now to make it happen sooner, rather than later.	

C O PAPERS I G N 2 0 0 1

New media semiotics – computation and aesthetic function

Craig Lindley, Interactive Institute, Sweden craig.lindley@interactiveinstitute.se Frank Nack, CWI, The Netherlands Frank.Nack@cwi.nl Andy Clarke, Kinonet Consultancy, UK andy@kinonet.com Grethe Mitchell, University of East London and Kinonet Consultancy, UK grethe@uel.ac.uk , grethe@kinonet.com Clive Fencott, University of Teeside, UK P.C.Fencott@tees.ac.uk

COSIGN 2001 is the first in what we hope to be an ongoing series of conferences investigating various aspects of the meaning, understanding, generation, and function of what may generally be referred to as new media. By new media we mean media forms that depend intrinsically upon computational processes to convey and realize their meaning. The papers selected for the conference have been chosen to cover a broad range of topics and issues from the perspective of media analysis, media production, and technology development, with the aim of encouraging the development of cross-disciplinary perspectives. This is crucial, since computer systems are essentially a metamedium, and every particular implementation of a software platform supporting a model of a new media artefact is in effect defining a specific medium. Hence there is not a clear boundary between software development for new media systems, and new media content creation. The interactive role of the media audience is also more pro-active than in traditional media forms, in many cases again blurring the distinction between the media producer and the media consumer. The result is the development of systems in which we are creating new forms of meaning, new modes of expression, and potentially new forms of aesthetic function. The diversity and unlimited potential of new forms complicates the analysis of new media systems. Understanding these systems therefore requires approaches that may identify principles of semiosis from a perspective from which form is fluid and highly variable. Within this problematic some themes emerge as highly general ones, such as the nature and function of narrative, a topic at the heart of most of the papers of the conference.

In this paper we present our own analytical narrative path through some of the topics and themes addresses by the papers in

First published at CO SIGN-2001, 10-12 September 2001, CW I,

 $The \ N \ etherland \ s.$

the conference. We have tried to bring out some of the issues that we regard as being critical ones in the development of our understanding of new media systems, and to interrelate aspects of the treatment of these various issues by different authors. The resulting coherence of the story establishes the validity of computational semiotics for new media as a coherent field of enquiry, and reassures us that a solid body of theory can be developed to provide strong and durable foundations for the field. Of course, many other theorists, media practitioners, computer scientists and artists are addressing these issues within their own frames of reference. We hope that our contribution to this process is to provide an interdisciplinary forum that can cross some of the cultural boundaries that might otherwise obscure or inhibit development of the field.

In "Bridging the Semantic Gap in Content Management Systems: Computational Media Aesthetics", Dorai and Venkatesh propose the use of high level semantic descriptions in support of search and retrieval of time-based media databases. Over the last few years, there has been a great deal of research addressing the automated characterisation, description, and annotation of image and video data content, in support of search, retrieval and browsing operations modelled on the older technology of textbased information retrieval systems. Having it's theoretical origins in signal processing, digital media content analysis has been focussed on media features that are comparatively easy to characterise, using descriptions such as colour histograms and mathematical texture measures. The rate of change of such lowlevel data characteristics has been used with a high degree of success to infer information about camera operations within shots, and various different shot transitions in digital video data. A long tradition in image understanding research has also provided techniques for detecting objects within images. Mosaicing allows the construction of large backgrounds by combining the shifting background contents of moving or different shots within a single location, and relative motion information for regions within images can be used as a foundation for the separation of background and foreground objects, potentially for reuse in different images. More complex situations, events, and scenarios within still image and video data have been regarded as a form of higher-level semantics to which content analysis research aspires in the longer term, but has had little or no success with in the shorter term.

This kind of high-level semantic content extends through more culturally specific connotations, to very specialised or even idiosyncratic interpretations of subtext and implicit significance. Research in content analysis has shown little awareness of these more specialised forms of meaning, although these meanings are frequently a very conscious object of intense exploration for media creators, critics, and cultural theorists. Dorai and Venkatesh propose the development of a more rigorous account of higher-level semantics, generally discussed in their paper in terms of editing principles and rules of cinematic style. This account, referred to as *computational media aesthetics*, aims to capture the higher-level semantics of user queries, which can then be related to lower level and readily computable data features. The examples explored in detail in the paper include the use of computable measures of pace, cutting rhythm, and movement rhythm, that are correlated with story segments and events at a higher, although generic, level.

This project amounts to the definition of a semiotics of computable features, in finding mappings between these features and higher-level descriptions expressing the semantics of dramatic structure. The feature sets thereby become signs, signifying the dramatic affective potential of video data segments.

The immediate contribution of semiotics to this project lies in the provision of existing distinctions that can be used to classify dramatic and cinematic form and function. However, consideration of the project raises question of the nature of a media database as a media form. Traditional information retrieval (or IR) systems are computational tools facilitating access to digitised collections of textual media artefacts. The paradigm is the well-established one of the database, in this case typically indexed by statistical content feature measures, and accessed via a statistical matching process against the text terms of a user query. Use of such a tool provides a particular view, and hence particular patterns of reading (or viewing, for a video database). These patterns differ from the linear, encyclopaedic, or random access structure of the media data units within the database. The IR model is therefore a media form in its own right, although the IR form as an aesthetic is highly constrained by its underlying models of search and browsing. The application of the IR model to traditional linear video in order to support search and retrieval (by film researchers and archivists, for example) is generally regarded from a much more instrumentalist viewpoint. It can indeed be regarded as a paradigmatic manifestation of Marshal McLuhan's observation, that when a new medium emerges, it is typically used to solve old problems, perhaps until a generation has passed. It is only after that initial period that the new medium starts to be applied to modes of creation and manifestations of aesthetic form that are uniquely its own.

This question, of the exploration of digital media as a new aesthetic medium, is at the heart of the concerns of this conference. If we take this concern into the realm of information retrieval systems (whether of the monolithic database kind, or the hyperlinked www kind), emphasis must be placed upon the

linkage and indexing systems as authored artefacts and crucial elements of the (inter-)active text. It is this that must constitute the primary computational aspect of link-based media systems, and hence a core topic within the development of a computational media aesthetics. In particular, linkages must be regarded as part of the text. This is highlighted in the paper by Kriechbaum and Stenzel ("The Link is the Data"), in which the relationship between the analytical representation of a media artefact (eg. a music score) is related to its physically derived representation in the form of the media data (eg. a sound file). Even in the case of more abstract analytical interlinkages, the relativity of interpretation holds, such as the relationship between the ideal architecture of a sonata form and specific sonatas that depart from that form, which might or might not then be regarded as erroneous.

Applying traditional media models to hyperlinked media forms generally meets with very limited success. This issue is explored by Fagerjord, who compares the structures of three print articles, three television documentaries, and three web sites, all of which are documentaries. Fagerjord finds the hypertextual web systems to be far more linear in their organization than hypertext theory suggests that they should be. In fact, the web documentaries fall into a common encyclopaedic model, a model of ergodic text (explored by Aarseth) for which computational navigation adds little to precomputational text models. More than a crossreferences structure is required to turn ergodic text into a new medium.

Srinivasan et al's Photoglass system more explicitly exploits the link structures of the www to create a new media model. Photoglass is a system that aims to create more realistic understanding of news topics by the juxtaposition of photographs from news stories from different sources. The photographs are selected and associated with topics by matching sets of keywords associated with the photographs and with a set of topics.

Photoglass again raises issues in need of a closer hermeneutic analysis. For example, the photographs, news stories, keyword lists, topics, and the links between all of these objects, are all, typically separately, authored artefacts (someone takes a photograph, someone chooses a keyword, someone makes a link). Hence, rather than presenting the truth of a story, the system presents a complex and dynamic text having multiply compounded problems of hermeneutic analysis and interpretation. This problem applies to any system that synthesizes media experiences (eg. the spherical photocollage of Photoglass) based upon textual representations (eg. lists of keywords or topics). All aspects of these systems must be regarded as parts of the textual artefact (Aarseth's concept of Cybertext). The World Wide Web as a whole is such a text, and our pathways through it are both facilitated and constrained by the keyword and topic lists processed by search engines, typically using search algorithms derived from IR research. As we move from diegetic meanings (ie. object and world denotations), through connotations, to subtext, we move from highly normative interpretations of textual meaning to increasingly specialised, and hence atypical, even idiosyncratic, interpretations. Topic and keyword lists are usually highly normative, so our pathways through the text of the www represent a collective recycling through highly conventional, surface, and therefore less

informative meanings. The www becomes a codification of mass culture, and surfing the web becomes a ritual of (re-)affirmation.

In "In Search of the Lost Text: Literary Theory and Computer Games", Kücklich more directly addresses these issues in the case of interactive computer games. Kücklich regards the program code, the set of rules governing the fictional world of a computer game, as the basic text (the media artefact), while the game itself is a playing, or reading, of the text. Kücklich cites Friedman's argument that learning and winning a computer game is a process of demystification in which the player learns how to decypher the signs on the interface as manifestations of the rules determined by the game code; this constitutes a reversal of the reading process. Moreover, metafictional signals within computer games challenge the reader to undertake a subversive reading strategy whereby the rules, code, etc. can be manipulated. Kücklich endorses Aarseth's suggestion that the sense of narrativity gained from a cybertext is a function of rapture and immersion, and that immersion within the narrative diminishes the critical distance required to develop a strong sense of story and plot. This issue of the relationship between narrative and interaction is central, and receives attention in many of the papers within the conference. It is also of general concern to the games industry, which would like to reach broader (and hence larger) audiences by achieving deeper levels of narrative: Kücklich's analysis suggests that this may not be possible, although it would be very premature to endorse this view. The issue is further complicated, not only by the ambiguous use of the term narrative by different authors, but by the ambiguous use of the notion of a game. The reverse reading strategy investigated by Kücklich matches the experience of game play as a decoding process for many common game genres, such as action games and strategy games. However, the model does not so easily fit interactive virtual environments in which there is no clear concept of winning. These environments, such as shared virtual worlds and massively multiplayer on-line role playing games (MMORPGs), can have an ongoing existence as cyberspaces within which the experience is one of interaction and world creation, not a decoding process but a process of ongoing encoding. Such a system provides a context for shared cultural articulation, having semiotics fundamentally different from game play.

When a virtual environment becomes a place for sharing experiences between more than one person, the semiotics of the experience can be expected to subsume many other, previously existing semiotic systems for interpersonal communication. Moreover, semiotic theories can also be applied to the analysis (and synthesis) of interactions within multi-agent systems that include both human and computational agents, as argued by Petric et al ("A Missing Link: The Role of Semiotics in Multiagent Environments"). This principle is further developed by Tronstad ("Semiotic and nonsemiotic MUD performance") in exploring the distinction between theatrical (semiotic) and (non-semiotic) verbs within Multi-User performative Dungeons/Dimensions/etc. (MUDs). Tronstad equates the theatrical with interplayer communication, affecting the player more than the character or state of the MUD world. The performative is (inter-)action that effects a change of state within the MUD world. However, the issue is ambiguous, since emotive verbs can have performative effects upon players. This difficulty

in the analysis highlights the simultaneous operation of multiple semiotic systems, or multiple simultaneous processes of semiosis; the theatrical ve performative distinction can be regarded from the viewpoint of player-to-player interaction, or from the viewpoint of the MUD as both a medium for that interaction and a shared artefact of creation as the subject of player intercommunication.

Tronstad also considers the nature of the quest from the perspective of this distinction. Tronstad regards a quest as primarily performative; a quest promises meaning, and finding the meaning (the object of the quest) kills the quest. At that point, once the quest has been solved, it retrospectively becomes a story. This suggests that, from the authoring perspective, the text (in Kücklich's sense) defines a specification of possible events, while the quest functions to drive the user through the available choices to create a game experience as a specific path through that space of possibilities; after the experience, the path can retrospectively be regarded as a story. Skov and Anderson ("Designing Interactive Narratives") address the issue of world authorship from a similar perspective. Basing their analysis upon interviews with two interactive storywriters, they suggest that the author is a world builder, not a storyteller, and that the story is created by user interaction. The question then becomes one of how it might be possible to characterize those worlds that are particularly good for generating stories. At least, this is suggested as an approach to interactive story building, since Skov and Anderson also observe that most games actually only allow users to vary the plot, while the underlying story does not change. In the interviews conducted by Skov and Anderson, three specific narrative structures emerge: the single selection path, multiple selection paths, and multiple exploration paths. Skov and Anderson suggest that the multiple exploration path approach corresponds to world specification, and go on to propose the use of thematic analysis to provide design patterns for interactive narrative systems; a thematic pattern could potentially provide the specification of narrative possibilities within a virtual world.

Despite a widespread concern with story in interactive systems, Julian Oliver ("Polygon Destinies") suggests that, at least for computer roleplaying games (RPGs), story is merely a context for the more important function of gameplay. Oliver observes that most RPG players are motivated firstly to explore, and only secondly to make it through to the end (to complete the quest). Again, story in current systems is generally not interactive. Each scene within an RPG is constructed as a means to the next scene, so RPG events feel and operate like a pre-eminent structure or fateline giving a particular significance to being in a particular place, since place corresponds with a position within the unfolding story. The importance of place creates an intensity unavailable just by being in a position within a story. The sense of intensity is also enhanced by objects, situations, and chains of events within the RPG world that work in themselves and together to produce learnable operating systems that transcend individual games, and even genres. The richness of the functioning, operational system reduces the need for imagination in order to enter the fictional world, allowing the user to more easily forget its fictional status. Oliver notes that the avatar in an RPG is not a visitor to the game world, but a denizen of it; its body and capacities are a direct expression of its virtual environment. The third person point of view typical of RPGs

enhances projective empathy with and sense of responsibility for the avatar, producing a kind of cybernetic selfhood.

Here again we encounter the significant difference between the decoding process required to win computer games noted by Kücklich, and the process of learning the operational system (and hence codes) of a game in order to enter the game world noted by Oliver. The RPG player learns the codes of the world in order primarily to enter and explore it. We can add that the quest, as noted by Tronstad, can intensify the process of exploration, adding an overall meaning and purpose to it that allows for mythical transformation into a retrospective story once the world has been finished. In fact, the quest is crucial for any concept of having finished the world, other than through pure boredom with endless exploration, or re-traversal of familiar territories. Indeed, it is the quest that transforms the world into a game, and creates a criterion for its completion. The game has criteria and required performances for winning, and it is the discovery of those criteria and the execution of those performances that constitutes Kücklich's deconstruction of the game rules. So here we have two rule sets, one needed for entering into and exploring the world, and another needed to complete its quests and finish with (or teleologically exit) the world. The world as a place can be a medium for interplayer communication (Tronstad's semiosis), and it's use as such a medium requires learning its operational system. The introduction of the quest requires learning the completion codes of the game, and moving through the game states to discover the end of the quest, constituting a performative process in Tronstad's terms.

Game and place can both be regarded as codes for the creation of narrative in its broadest sense, as a coherence of experience. However, other forms of coherence are possible, and incoherence can also be a goal for various forms of aesthetic function (including the creation of incoherence according to normative semantics, for the sake of driving interpretation to levels of subtext for the discovery of less normative forms of coherence). In general, terms like coherence, chaos and order, cannot be allowed to go unchallenged. In "Introducing Interrelations of Madness and Virtuality", Walz and Zundel consider the interrelationship of the codes and mappings respectively of madness and virtual environments, asking what understandings one can give to the other. In particular, it is suggested that nonanthropomorphic works in a CAVE (an immersive, 3 dimensional virtual environment) constitute a virtual symbolic chaos in relation to the symbolic order of (non-virtual) reality that is analogous to the symbolic chaos involved in representations of madness in which madness is opposed to the symbolic order of sanity. The notions of chaos and randomness used in this discussion do not conform to their more technical, mathematical definitions, since chaos is not the same as randomness within complex dynamics. Rather, chaos is about order within complexity, and randomness is about absence of information. So perhaps the more interesting issue here is to leave behind folk conceptions of madness, and explore the richer areas opened up by applying less dismissive interpretations of madness to the interpretation of VEs (anthropomorphic or not). That is, within semiotics, the notion of madness as symbolic chaos falls apart under a psychodynamic view, since the program of psychoanalytical interpretation is that of creating an account of the underlying systems and structures responsible for

normatively unpredictable and/or socially dysfunctional modes of discourse historically and culturally designated by the term madness. This method of analysis can be applied to the experiences with a VE, whether or not those experiences are anthropomorphic, and irrespectively of any categorizations in terms of madness or sanity, the latter holding little interest or information. Interpretation from a psychoanalytical perspective is well within the traditions of semiotic analysis.

Turning this around, to synthesise the contents of a VE dynamically and interactively, based upon an underlying representation expressed and understood in the language of psychoanalysis is an interesting area for the ongoing development of VEs. The development of VEs based upon the mathematical fields of chaos and complexity is also a very interesting area for ongoing research. And we can ask how these two might be brought together, for example, by generating a chaotic virtual environment state space from a symbolic psychodynamic model (and what the relationship between the symbolic representation and the chaotic space can be), or to create virtual experiences having psychodynamic patterns or meanings, conceived of as strange attractors from a chaotic generative basis.

The creation of specific new media productions generally requires new media producers to grapple with many of the same issues as more theoretical analysts. In "Beyond Manzanar", Thiele describes the creation of an interactive virtual reality (VR) system having a spatial and transitional design supporting user experiences following a traditional dramatic structure, from build-up, through a climax, to a denouement. This amounts to the creation of a space of predetermined narrative possibilities, following the underlying structure of the dramatic dynamic. Nakai et al describe the Cypher system that allows people to create virtual spaces by the manipulation of real 3D blocks in physical space. The users of Cypher can then create a photograph of themselves within the 3D world that they have created. Composition of the photograph is determined automatically by comparing the user's pose with a database of "masterpiece" paintings, and replicating the position of the figure in the field of the most closely matched painting.

The Cypher system is another instance of the use of a new medium to solve a problem arising from an old medium, that is, the problem of how to create a photograph of oneself in an unreal setting. The result is to be able to represent yourself to others as a participant within a virtual environment that you have never actually experienced as being within. An interesting question here is that of the nature of such artefacts, not as real photographs, but as possible virtual artefacts within VEs. A virtual photograph might never be represented outside of a VE. We might then ask how the nature of a photograph is transformed by its virtuality. That is, what does such a photograph become, when every aspect of it is potentially (endlessly) modifiable? Instead of being a fixed death mask, the photograph becomes a kind of playing card, a 2D form within a 3D world, carrying an image and function that is always open to suspicion. The image has an unknown history, ambiguous authenticity (as what?), and could change in any way at any time. Moreover, the virtual photograph could also be any number of virtual interface devices, with unknown functions in unpredictable areas of the virtual world. What the photograph

signifies has therefore radically changed, into an inversion of its traditionally conceived nature as an authentic and fixed image in the real world.

The redefinition of familiar forms within virtual environments opens up endless possibilities for the creation of new signifying languages. An example is the system described by Fuchs and Eckermann ("From First Person Shooter to Multi-User Knowledge Spaces"), a VR system for cross-disciplinary virtual museum exhibition. In this system, architectural form is used to convey the semantics of links and exhibit objects. This is an extension of traditional semiotic functions of architectural form to allow the user's exploration of a quasi-spatial structure to function as the key mechanism for creating a semantic structure that is neither linear nor hierarchical. Hence the space becomes the semantic index, and the user can go where they like, allowing for individually shaped relational networks inside a complex field of knowledge. Fuchs and Eckermann ask: given a clear semantic structure, then mapping semantic fields to rooms, signifieds to topics, and signifiers to 3D objects and sounds, does the game the user is immersed in still correspond with the semantic structure? Their answer is no, due to technical limitations, subjective connotations, etc.. However, this is a very interesting general question, of the effectiveness of a symbolic, semiotic use of spatial design that requires the establishment of language-like conventions for people to be able to understand the (in many cases) arbitrary mapping of spatial structure to other forms of meaning. This is a complex issue of leaning new meanings, since the symbols already carry meanings that must be distinguished from the supplementary language of spatial symbolism that could include many specialized and new associations. However, such languages are certainly feasible, potentially constituting the form of learnable operating system discussed by Oliver (see above).

This creation of new semiotic systems is a common process within new media creation. Another example is Seine hohle Form (Rovan, Wechsler, and Weiss), in which dance gesture is used to drive sound synthesis. Superficially, ShF could be regarded as the definition of a signification system, and in fact a symbolic language, for mapping semantically arbitrary gestures onto synthesised sounds. However, the gestural language of dance is in this case a semi-autonomous aesthetic system, the aim being to fuse dance and audio performance at a more fundamental level. The result is an audio, visual, and kinetic system that functions as a synaesthetic whole, affected as a communicative system acting between the performer(s) and the audience. The creation of mappings between gesture and sound hence constitutes development corresponding to the phonetic and phonemic level for a genuinely new and hybrid language.

The issue of the creation of new languages versus the use of newly created languages raises an important issue in the generation of new media systems. This is the issue of where the boundary lies between being a system builder and a media artefact author. The issue is important, since it relates to the commercial domination of new media, and the need for content production to be accessible while not detracting from the depth and breadth of media productions. For example, Portlock ("Streams of Motion (Super Spectacular) A Virtual Reality Art Work") presents a personal case study of creating a VR artwork as a single person production. The question that arises is that of how a rich and complex system can be created with very limited resources. There have been very few non-commercial artistic explorations involving rich, complex virtual environments, and this is due to the large resources required both to undertake the programming tasks involved, and to produce 2D, 3D and audio content for those works. Budgets and teams for producing commercial computer games are increasing to the level of commercial cinema productions, resulting in a mass market aesthetic for these games aimed at recouping their high production costs. The result of this is visible in the highly conventional forms and genres of most computer games. Better tools are needed to allow people with limited resources to produce rich and complex virtual worlds, for the sake of creating more diverse types of worlds of interest to more limited audiences, and to explore many forms of aesthetic function outside the interests of the commercial media industries. An approach to creating such tools is to look at more generative approaches in which media content is created, formed, or grown at a high level, or via interaction within a virtual world. Such systems begin to blur the boundaries between media tool/systems producers, and media content producers. In fact, they begin to blur the boundaries between media producers and media consumers. In these systems, the languages that define the spaces of possible content allow media producers and consumers to explore those spaces by interacting within virtual worlds. The authorship of the resulting experiences is then highly dispersed, amounting to collaboration at all levels. This is a fascinating area for further exploration, theoretical study, research and development, although it is not an area that has been addressed within the current set of papers for the conference.

"The Data][h!][bleeding T.ex][e][ts_" by Mez Breeze presents an appropriate finale for this discussion of new languages. "The Data][h!][bleeding T.ex][e][ts_" is an eruption of new language, in this case reflecting a life immersed in hypertextual structures and nonlinear media. In Breeze's work, conventional linear writing is fragmented and interpenetrated with associations and fragments of associative text and the lingo of the worlds of the Internet and web art, email, MOOs, and a cyberfeminist joy in disruption and transgression by the electronic word. This is hypertextual writing liberated from the rationalist grids of encyclopaedic organization into a poetic transtextuality, a scriptural splitting even within the structures of individual words. Perhaps here we see the symbolic chaos referred to by Walz and Zundel, in the schizoid preoccupation with cybertext as a field of textual play, and the revelling in foreground, surface, and association that free the cyberwriter from preoccupations with narrative, creating instead a form of engagement and immersion in which the self is lost (no more heroes).

It's intriguing to wonder what the 3D visual analog of this associative style might be. Perhaps this would be the ultimate generative medium, in which the hero's journey is abandoned for a visionary field of synaesthetic forms, unfolding in fractal patterns and symbolic associations. If we add to this the integration of unconscious biofeedback (eg. see the poster by McMahan), we can remove the need for conscious intervention in unfolding experience. The collective information space then becomes a free-flowing pool of meanings, from the idiosyncratic to the archetypal. This is the other pole of the universal computational(ly facilitated) mind, not the super-intelligent rational machine, but the collective realm of resonances and patterns, mystifiable under the title of the universal subconscious, dream machine of the species.

Issues of Sequence in Converging Media: Studying World Wide Web documentaries.

Anders Fagerjord

Department of Media and Communication, University of Oslo. P.O. Box 1093 Blindern 0273 Oslo, Norway Tel. +47 22 85 04 11 anders.fagerjord@media.uio.no

ABSTRACT

The paper reports on a study where three print articles, three television documentaries, and three Web sites by the *National Geographic Society* were compared, to see what constitutes linearity, and whether nonlinearity takes the place of linearity in Web media.Devices of transition, disposition, and linking are studied, as well as aspects of narration. The Web sites proved to be linked in a manner that most of the time only allows a linear reading.

KEYWORDS

Web, film, print, linearity, nonlinearity, multicursality, sequence, transitions, links, illustrations, semiotics, narratology.

1. INTRODUCTION

Studies of computer texts (using 'text' in the wide sense) tend, in varying degrees, to point to both the novelty of computer texts, and their dependence of earlier media. While keeping a historical perspective, books like Jay David Bolter's *Writing Space* [9], George P. Landow's *Hypertext* [26], and Espen Aarseth's *Cybertext* [2] attempt to set computer texts apart from earlier forms of writing. More recent books, such as Bolter and Richard Grusin's *Remediation* [10] and Lev Manovich's *The Language of New Media* [30], explain digital media from their dependence of earlier media, such as painting, television, and cinema.

Studying Web media, I believe we have to take both approaches. Web pages can combine semiotic systems: text, recorded sound, still and moving images. We have only so many codes for using these different semiotic systems, and even fewer for how to combine them. These codes, or rhetorical practices, are heritages from earlier media. Then we have the new possibilities of the computer: the ability to mix these semiotic systems (multimedia), the global two-way network of the Internet, the speed and programmability of computers, and the flexibility of digital text and images, and so on. Such new possibilities are increasingly influencing older ways of constructing messages. I believe it useful to view Web sites as instances of rhetorical convergence: products of technological con-

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands. vergence and in itself a convergence of media practices — rhetorics — formerly found in different media, analog and digital. What are these semiotic codes and rhetorical practises that are converging? They are conventions, ways of constructing texts proven to be effective in the different media. Techniques such as continuity editing of film or the newspaper's "inverted pyramid" style cannot be deduced from the medium's technology alone. Moreover, many of them are mutually exclusive, they cannot be combined. Authors in a convergent multi-medium have to choose which ones to use, at the same time excluding other possibilities.

In order understand how this determines the rhetoric of popular Web sites of today, we must study actual sites, and compare them to "old media". The most theorised aspect of hypermedia is probably their alleged nonlinearity, and this concept seems to be a good place to start probing into the semiotic of Web media. Nonlinearity is normally defined through what it is not: linearity. But is linearity well enough understood? Only by knowing the difference between the two can we hope to understand nonlinearity.

In this paper, we will examine how print and television documentary construct linear readings or viewings. Then we will see whether parts of this linearity can be found again in reading the Web, or whether multicursality (non-sequentiality or nonlinearity in other vocabularies) takes its place. If we find multicursality, we will study what the mechanisms of this multicursality are.

2. NATIONAL GEOGRAPHIC IN 3 MEDIA

The basis for this paper is a study of a selection of Web "features" at Nationalgeographic.com, television documentaries on *National Geographic Channel* (Europe), and illustrated articles in *National Geographic Magazine*. They covered three themes: (1) Meteorites: a Web site, a television documentary and a magazine article about the danger of meteorites to the earth; (2) Volcano: a Web site, a television documentary and a magazine article about the volcano eruption of the Caribbean island Montserrat in 1994–96; (3) Midway: a Web site, a television documentary and a magazine article about the Battle of Midway in the North Pacific during World War II and the 1998 search for the sunken aircraft carriers from the battle. In total, nine texts were thus analysed [1, 3, 11, 21, 22, 31, 36-38]. All three web sites in the selection contained video, making comparisons to both print and television natural.

The texts (including the films and Web sites) are in a genre we could label science communication, a sub-genre of feature journalism or documentary film. They are written to communicate knowledge on geography, culture, zoology, the environment and other sciences. The National Geographic Society also has a tradition of sponsoring and writing about daring expeditions and explorations to remote parts of the World. These different kinds of knowledge are presented in ways believed to be clear and understandable, and at the same time entertaining and exciting.

To achieve these ends, print journalists rely on the devices known from other kinds of feature journalism They base their stories on interviews and compelling description interspersed with more expository paragraphs. Documentary film also use interviews as a basic technique, between parts in the classic documentary mode with the "camera as eyewitness" often featuring a narrator's voiceover to explain. A widespread practice within documentary and print journalism is to try to "tell a story", which usually means to pursue involvement through the reader's identification with persons in the text and a dramatic curve. Often, this is done by making the text a narrative.

Thus, all the texts in our sample are blends of three modes traditionally labelled Narrative, Description, and Exposition. Seymour Chatman calls these modes text-types [13], and argues that they are hard to discern in practice. These text-types might not hold against a fine-tuned semiotic analysis, but in this paper they serve the purpose as descriptions of three different modes of communication. We will see that though they are covering the same themes, the Web sites, articles and films still are structured very differently. This should not come as a surprise, as there always will be a number of different ways of structuring any text.

3. A SEQUENCE OF PARTS

Many of the world's printed texts are written to be read from top to bottom. Most speeches are delivered in one sequence. Television programs and movies in theatres also have one sequence they are viewed in, from beginning to end. These are the texts we will call linear. Writers have always known that to write a good text, it is important to make this sequence with care, as several sequences are possible from the same material, some more effective than others. In Aristotle's words: "There are three things which require special attention in regard to speech: first, the sources of proofs; secondly, style; and thirdly, the arrangement of the parts of the speech" Rhet. 1403b[17]. Different kinds of texts fulfil different functions, and have different parts and sequences. Again, this has been evident at least since Aristotle discussed proper sequences in his treaties on the kinds of speech and of drama in *Rhetoric* and *Poetic* respectively.

In my analysis of National Geographic material, the texts, films and Web sites were divided into parts. This is done by finding ruptures in the text, points of weakening coherence in Michael A. K. Halliday's terms [19]. Although our texts all are coherent, the different parts establish their own coherence by local cohesives in Halliday and Hasan's terms: references to persons, objects, environments or arguments not found in the parts before and after. As such, a part is an internally united unit that stands apart from what proceeds and follows. In practice, in some parts of the material analysed, the parts were easy to establish, while other sections were harder to divide decisively. This should be expected, texts and films are intricately woven fabrics of signs, and this difficulty just demonstrates the text's overall coherence. After distinguishing the parts in each media text, their order was analysed. Some media-specific differences in between articles, films and Web sites can be characterised using the classic distinction of narrative, description, argument, and exposition. The three television films in our sample are narratives, following a limited number of characters through dramatic events. Less unified are the three magazine articles, Midway being narrative, Volcano containing equal amounts of narrative and descriptive discourse, and Asteroid being expository or descriptive. Of the Web sites, only Midway contains narrative. The bulk of content of the three Web sites is descriptive or expository.

Whether narrative or exposition, the text-type does not automatically command an order for the text's parts. But there usually is logic to be found, at least in mainstream journalistic discourse like our material. In my (linear) magazine articles and TV documentaries, I found three different principles of establishing a sequence in the linear media:

- 1. Chronology (*Volcano* film, *Midway* film, *Midway* article, Asteroids film): Although containing achronies like flashbacks and flashforwards, a narrative progresses in time.
- 2. Geography (*Volcano* article): The order of the parts follows a line in space.
- 3. Classification (*Asteroids* article): The article explains about different kinds of meteors, historical accounts of such, their origin in space, ways of finding them, and kinds of effects of impact on earth.

Chronology structures four of the six texts. This is not a simple chronological sequence, as all the texts contain several flashbacks and flashforwards, what Genette calls anachronies [18]. In fact, a plotting of these anachronies the way Genette does in *Narrative Discourse*, show structures just as complex as many of Genette's examples from Proust. For example, the *Asteroid* movie starts out simultaneously in 1997 (images) and the 1940's (voice-over), moves into a hypothetical future, then to March 1993, 1995, pre-1950, 1997, future, and 1947 in just the first six minutes. Still the overall movement in the film is from ca. 1945 to 1997.

Although listed as chronologically ordered, the *Volcano* article is a narrative, relating the story of the writer's visit to the island Montserrat. The story follows her travels around the island, telling the experiences of the people she meets in each place.

The Asteroid article is classificational in structure, that is, it sets up a taxonomy of different kinds of extra-terrestrial objects hitting the earth, the biggest, the smallest, the earliest described, etc. It moves from general to more detailed descriptions, building on what is already said when presenting detailed material. However, this is only a loose ordering structure, and other devices can also be found. It opens with a description of a scene of a huge meteor impact 50 000 years ago, inserted in the middle is the story of a comet impact in 1908, and it ends with the threat of human extinction. These three dramatic stories help establish a dramatic curve. The flow of the article is also helped by smaller-scale techniques, what I call transitions.

3.1 Tying the Sequence

3.1.1 Transitions

Although the discourses in our linear media sample consist of separate parts, they are felt as a continuous line, constantly working to keep the interest of the audience. Smaller-scale techniques are put to use in the very transitions from one part to another. Handbooks for writers offer examples of such transitions. In one such book, *The Random House Handbook*, Crews lists 11 kinds of "signals of transition" that "indicate that the previous statement will be expanded, supported, or qualified in some way": Consequence, likeness, contrast, amplification, example, concession, insistence, sequence, restatement, recapitulation, and time or place [14, p.113].

In my material, I have found 10 transitional devices, greasing the movement of linear texts.

- 1. Chronology ("later").
- 2. Reversed chronology ("before this").
- 3. From particular to general (metonymy).
- 4. From the general to the particular (example).
- 5. Causality (from cause to consequence).
- 6. Reversed causality (from consequence to cause).
- 7. Similarity.
- 8. Contrast.
- 9. Movement in space.
- 10. Logic (premises and conclusions).

These are significations of semantic relations, they spell out the relations between sememes. I found these 10 devices in both print and television. The TV documentaries analysed have even more transitional effects at its disposal. Television films can employ at least five systems of signification (tracks) simultaneously: images, text, spoken language, other sounds, and music. Not only can different transitional relations be used simultaneously in different tracks, but they may also be asynchronous. A cut in one track is less disruptive if one of the other tracks continues. This is a common editing practice, found in a lot of instances in the films. Other editing practices achieve similar effects, such as continuity in space and movement between shots, and dissolves between similar-looking objects, a sign relation that is unique to filmic media (Figure 1).

We kind find reasons for the use of these transitions in different semiotic theories. Both Umberto Eco [16] and Halliday [19] point to the need for predictability in communication. The addressee



Figure 1. Asteroid film: A dissolve between the result of a gunshot into sand (left) and Meteor Crater in Arizona (right) demonstrates the similarities.

needs to be able to guess fairly accurately what will be said next in order to comprehend easily. It follows that a rupture in a text, a change of topic, will be felt less abrupt if the author makes an effort to spell out how the new topic relates to what came before. This semiotic of transition, based on relations between sememes, on similarity between signifiers, and on breaking at different points in different semiotic systems, is a gloss effacing a well-known fact: Although the sequence of a text's parts is determined (e.g., by narrative or argument), the sequence is only one of a large—perhaps unlimited—number of equally possible sequences.

The ten transitions listed above can be found in all of the articles and films analysed, regardless of their organisation (chronology, geography, or classification) or dominant text-type (narrative, description, exposition). As noted, all but one of the articles and films are dominantly narrative, and this powerful device has its own ways of uniting a line.

3.1.2 Narrative Ties

Roland Barthes wrote about linearity in S/Z. His analysis concludes that linearity is dictated not by chronology as such, but by the sequence of actions, and the control of knowledge in the discourse, what he named the hermeneutic code.

The five codes mentioned, frequently heard simultaneously, in fact endow the text with a kind of plural quality (the text is actually polyphonic), but of the five codes, only three establish permutable, reversible connections, outside the constraint of time (the semic, cultural, and symbolic codes); the other two impose their terms according to an irreversible order (the hermeneutic and proairetic codes). The classic text, therefore, is actually tabular (and not linear), but its tabularity is vectorized, it follows a logico-temporal order. It is a multivalent but incompletely reversible system. [4, p. 30]

These codes are distributed all over the text, and give coherence to the storyline as a whole. Barthes's five codes were formulated to fit a piece of classical narrative fiction, and not factual texts, whether narrative, argumentative, descriptive or expository. Thus, they may not fit our material exactly, but I believe they do help us understand how the narrative texts succeed in establishing a storyline that keeps the interest of the reader.

For instance, Barthes put references to chronology as a subclass of the cultural code, as references to a chronology is a way of lending realism to a fictive story-world, an effet de réel [4, p. 273]. In documentary texts like ours, however, the chronology of events might actually be an important part of the story conveyed. Furthermore, not only giving fact or realism to the text, markers of chronology are important in understanding how the parts of a discourse should be pieced together into a story, as events can be told out of order. As long as the reader can recognise an event's time in relation to other events, she can easily put it into its place in her internal map of the story's chronology. This does not come out in Barthes's analyses, as the texts he analyses are perfectly chronological in order.

The linear stories have long chains of actions spanning large portions of and long periods of time as well as shorter ones. These long over-arching chains of actions keep linear texts together in one single line. In the Asteroids film, a running theme throughout is geologist Eugene Shoemaker's struggle to win acceptance for his theories. Again and again, the film refers back to this theme, adding a new link in the chain of events, bringing coherence to the film.

As seen in the quote above, Barthes believes that in addition to the proairetic code, or the sequence of actions, the hermeneutic code controls linearity in a text. The hermeneutic code is the control of knowledge in the discourse. Knowledge of what has happened; how and why it happened and with what consequences; when it happened; and how it all ended. Posing an enigma at the start of a story, and slowly releasing the answer keep the reader or viewer's interest. In the Midway film, we wonder until of the film whether Ballard will actually find the sunken *U.S.S. Yorktown*, while the narrator keeps returning to this question over and over.

3.2 Pockets of narrative

Both magazine articles and documentary films relay heavily on citing small descriptive episodes, small pieces of narrative: a trip, a visit and so forth.

Chatman [13] points out that narrative and description may be found in texts that are dominantly argumentative or expository as well as narrative or descriptive. This is certainly the case with our material, where these episodes are found in all kinds of texts. We will call them episodes as they usually are quite short, and may be about a person we will not meet again. In both kinds of text, they often stand as (metonymical) examples of a greater problem or a repeated task; showing, not telling.

Usually, introducing a new episode makes the transition to a new segment. The present episode is closed, by establishing an intermediate closure, and in the films, the camera usually moves away. Either it moves with a character's view to a distant object, or it tracks away from the scene and the character. A new time or place, other persons are brought upon stage, usually signalled by a conventional establishing shot or a shot of people arriving, while the narrator's voice anchors the pictures in time and place. This structure of arrival-scene-departure gives a special effect: For a film about asteroids, comets and meteorites, Asteroids have a remarkable number of shots of cars driving to places! The events in the film Return to Midway are mostly taking place in one spot in the North Pacific, but the editing achieves the same effect of intermediate closure by piecing together a cyclic movement from sea to air to sea again. Return to Midway mobilises an impressive amount of documentary footage from the 1942 battle. Each round in the World War II air battles start with planes getting ready and taking off, the camera being on the carrier ship's deck. Then the planes are filmed in the air, and bombs falling, the camera being on level with the planes. As the bombs fall, the camera's point of view shifts to ground level, where the effect of the impacts is shown, as well as planes crashing into the water. From sea to air and back to the sea, the cycle is complete, and the episode is connected and well rounded off visually.

We will return to these narrative devices when discussing narrative in the Web sites, but first, we will look for nonlinearity.

4. WITHOUT THE SEQUENCE

Many writers throughout history have related their trouble in organising their thoughts in a linear fashion. Gunnar Liestøl, for example, finds examples of such laments in the philosophies of Plato, Hegel, and Wittgenstein [29]. In the origins of hypertext we find a similar voiced dissatisfaction with the linear organisation of alphabetic writing. When Vannevar Bush launches the influential idea of the "memex" in his 1945 article "As We May Think", it is a library ordered by association, like the human mind, rather than hierarchy. In this library, any record can be part of a number of "trails"; the libraries have "a mesh of associational trails running through them" [12]. In 1965, Theodore H. Nelson writes: "Let me introduce the word 'hypertext' to mean a body of written or pictorial material interconnected in such a complex way that it could not conveniently be presented or represented on paper." And in a footnote he adds: "The criterion for this prefix (hyper) is the inability of these objects to be comprised sensibly into linear media [...]"[32]. He reformulated it in 1974: "By 'hypertext' I mean non-sequential writing. Ordinary writing is sequential [...]. But the structure of ideas is not sequential. They tie together every which-way [33]". In Tim Berners-Lee's original World Wide Web sketch of 1989, Bush's argument is echoed: "[...] a 'web' of notes with links (like references) between them is far more useful than a fixed hierarchical system [8]".

Scholarly studies of different kinds of hypertext have stressed hypertext's opposition to linearity, using the terms nonlinearity or multilinearity. Examples include Yankelovich, Meyrowitz and van Dam's "Reading and Writing the Electronic Book" [39]; Landow and Delaney's "Hypertext, Hypermedia and Literary Studies" [28]; John Slatin's "Reading Hypertext" [35], George P. Landow's "The Rhetoric of Hypermedia" [25], *Hypertext* [Landow, 1992 #23; as well as Espen Aarseth's "Nonlinearity and Literary Theory" and the 9 other essays in Landow's collection *Hyper/Text/Theory* [27]. Later, in Cybertext, Aarseth first replaces the term nonlinearity with multicursality, and then includes multicursal texts in his more nuanced concept of ergodic literature, literature where the user in a reading "will have effectuated a semiotic sequence", literature where "nontrivial effort is required to allow the reader to traverse the text", an effort beyond simply reading and turning pages.

5. LINKING

The device that potentially gives hypertexts the power to go beyond linearity is linking. I find it useful to discern between three kinds of links, operating according to different logics: Navigation links, presentation links, and relation links.

Navigation links are present on every page in a section of a site, most often on top of the page or in a menu bar on the left. They give the reader the ability to locate a part of the text through only a few clicks, thus their name. Navigation links give an overview of the whole hypertext by dividing it into categories. Hence, like sememes in a semiotic code, their logic is positional: a reader will deduce what probably is contained in a section not just by relating the signifier of the link anchor to its usual signified, but by considering what other categories there are. A section contains what is not in the other sections.



Figure 2. *Midway* Web site. Presentation links at the bottom of the screen, "file tab" navigation links at the top.

Presentation links are links used to display the next or previous page of a part of the text, to start a movie, and other controls of the display of the current text body. Typical presentation links are the "next page" button on the bottom of a page, or the "read whole story" links on the front page of a news site. They are similar to page turns in a book: necessary to proceed in the text, but not formative to the text's structure. In most books, most pages can be broken, if not at any place in the text, so at least a lot of places, without altering the text. The Bible is the same, whether it is printed on 800 or 1600 pages. Presentation links are essentially linear; the link text could be "proceed" (Figure 2).

The third kind of link is the relation link, which is a jump to another place in the hypertext that is related in some way or another to the present page or paragraph. Their logic is that of association. If the transitions we found in linear texts would have corresponding ways of linking, they would be transition links. But:

There are no relation links in the Web sites analysed, only navigation links and presentation links. As one would suspect, similar (or even identical) events, facts or information are treated in different pages of a Web "feature", but they are never directly linked to each other. The Midway web may serve as an example:

4.2 Linearity in the Midway Web.

In the *Midway* Web, Navigation links are found on the top of every page as "file tabs" (Figure. 2).

A lighter colour indicates the reader's present location. The tabs give access to the five main areas: *Ballard*, *Briefing*, *Ships*, *Dispatches*, *and Findings*. Readers are encouraged to read the areas in this sequence, not only from the ordering of them from left to right (or top to bottom on the first page), but also from links between bottom-level pages. Starting in the *Ballard* area, and clicking "next" at the bottom of each page, one is taken in an unbroken sequence through *Ballard*, *Briefing*, and *Ships*. We will call this the "default" sequence.

Readers do have some other options. Each of the five areas are subdivided, the subdivisions being listed in a column at left, and the lighter colour shows the location there as well. These navigation links give the user the possibility to read at a front-page level, by clicking the page-top tabs in any order, and to read at a skimming level, by clicking the buttons on the left. To read the full story of each subdivision, however, one has to read linearly, as the subsec-



Figure 3. "Catalogue card" metaphor in the Midway Web site.



Figure 4. *Midway* Web site. Clicking on a part of the ship in the drawing in the main screen loads an underwater picture of that section of the ship in the smaller window.



Figure 5. Illustration from *Midway* article combines classificational and analytical figures.

tions are further divided in to two or three pages each, with two "back" and "next" links at the bottom. It is possible to read the sequence of the first pages of each subsection backwards, but the reader will have to work hard to comprehend, as many pages need to be read in context of the preceding pages.

4.2 Description and Exposition in Web Media

If we as readers choose to follow the "default sequence" in the Midway Web, we move through different textual modes. After a question-and-answer interview in *Ballard*, a narrative starts: the story of the Battle of Midway. It is mainly chronologically ordered, and continues through 14 pages in *Briefing* and 4 pages in *Ships*. Then, the narrative "motor" comes to a halt as we come to the "catalogue card" section of *Ships* (Figure 3).

There are no more "next" links from this point on, but if we as readers continue to visit the navigation tabs in left-to-right order (the way we read in Western languages), we continue the chronological order: We have read why the story is told (Ballard interview) and the story of the Battle (Briefing and Ships), we continue with the story of the search (Dispatches), and the results (Findings). But when we state that the narrative motor stops in the Ships section, we do not merely talk of an absence of presentation links. The text is constructed in a completely different mode, it changes from narrative to description only. As the four catalogue cards with the ships are similar in content and layout, we are no more reading a story; we are comparing ships on a common scale. Readers get into a mode of research and comparison in this section; the hypertext merely offers its contents and tools to organise them. The pages are no longer organised to construct a sequence, but to present data in even-sized chunks, that the reader may group and regroup freely to gain overview. In the Findings section, the same effect is obtained when underwater photographs of details on the shipwreck are available from a drawing of the whole (Figure 4). This mode of standstill and comparison is equal to the illustrations Gunther Kress and Theo van Leeuwen [24] call classificational figures, which "relate participants in a 'kind of' relation, a taxonomy" (p.81) and analytical figures, which "relate participants in terms of a part-whole structure" (p.89). The illustration explaining the battle of Midway in the magazine article combines both forms (fig 5).

To read such a map, graph, or illustration is to perform comparisons in one's mind, according to the principle organising the signs, whether classificational or analytical. The signs are distributed in a space, their positions themselves becoming signs. But there are no sequence the signs present them in, as signs in sentences and paragraphs of written or spoken language or in film. The reader crisscrosses over the text, reading the signifiers and their positions in a searching pattern.

Our two Web examples consist of multiple pages, and do not, like images, present an all-encompassing overview. The reading process, however, resembles the reading of signs in a figure, essentially nonlinear or multicursal. The material is divided into parts that are comparable in scope and size, and each page is equally available through a set of navigation links that provide overview of the whole. We might say that these parts of Web sites are classificationally structured by navigation links. .Of the other two Web sites analysed, *Asteroid* makes use of a similar classificational link structure for its descriptive and expository content. Information on asteroids, meteors and comets is divided in equally structured pages, accessed from a list of links. In the *Volcano* Web site, however, an exposition of different kinds of Volcanoes is strictly linear, its pages connected with "next" and "previous" presentation links. The taxonomy presented is not represented either in a graphic overview or in the linking structure.

4.3 The Semiotics of Links

The structure of these sections gives the reader freedom to browse at her own will, thanks to the way the pages are linked and the way the links are signified. Links need to be doubly signified. Firstly, as links need to be activated, their presence is be signified. Link anchors can be parts of alphabetic texts, images, and parts of images. Normal codes for links are underlining, a different colour, or a different coloured background. Images can be signified as links by a coloured border, or by placement on the page, e.g., in the margins or next to a short text describing another page. Some sites make their own "icons", images that signify the presence of a link. A link can also be signified by a change when the cursor passes over the link anchor, by changing the link anchor sign, or by a change in the cursor symbol. I list these fairly obvious examples to point the many different codes that enable us to understand where the links are. Codes are not universally shared, few sites make use of all these codes, and many examples can be found of Web pages using a different colour or underlining to signify importance, devices that in other codes signify links. Frustrated users click on parts of text or images that they believe are links, before they search for the actual links using the cursor as an index: if the cursor symbol changes, the cursor is over a link.

Secondly, the destination of the link needs to be signified. Loading a new page is always a disruption, and it might be a severe disruption, displaying a page belonging to a whole other context. Above, we noted the need for addressees to form expectations of what will be communicated. If the reader is to form any expectations of what is to come when activating a link, the content of the destination page must be signified. This is done differently for the different kinds of links I have described. Presentation links does just have to signify that they are presentation links. They would typically read: "next", "previous", "read full story", "play video", or be a rightpointing arrow at the bottom of the page. The destination is thus implied: more of what the reader has already started. Placing links as navigation links, normally in the margins of the page, and often as a list of the particular Web site's sections, also signifies the page that will be loaded: the first page of the section indicated by the link. For relation links, there exist few such codes. Unless the link is within a sentence describing another page, a casual reader most often have a hard time guessing on which page he will land. Local codes exist in some sites. Most surfers know that when a name is linked, in some sites you can expect the link to bring you to the person's homepage or biography. Other sites use a code where a linked name can activate the user's e-mail client to open a new message to the person mentioned. In the Volcano Web, all links in the text open images in a small window. It takes a little reading to discover this local code. The link anchors are general words that in other sites would be relation links, opening other pages of text, so before the reader discover this local code, she might be surprised, even disappointed, that she is never taken anywhere else.

4.4 Multicursal Narrative

Of the three Web sites, only the *Midway* Web contains narrative. This is a marked difference from the films and articles, where narrative was the main text-type. Even the *Asteroids* article, mainly expository in mode, enlivens the discourse with several narrative episodes. In the *Midway* Web, we find no episodes. The narrative in the *Briefing* section tells the story of the Battle of Midway in strict chronology with no achronies, keeping a distant perspective to the events throughout.

We noted in the linear media that coherence was kept by distributing long chains of actions over the whole of the story. Also in this respect, the Web media analysed were simpler than print and video. Chains of actions are short, most often kept within a single page.

We have seen that the *Briefing* section of the *Midway* Web in places is more linear than a book: some pages in the Web are only accessible at the bottom of the preceding page, a traditional book can be opened on any page. Still, the subsections, collections of two to three pages may be accessed in any order. In the *Dispatches* section of the *Midway* Web, a story is told through 27 letters from a reporter. The 27 letters may be accessed in any order. This makes multicursal narrative a possibility.

4.2.1 Enigmas

In the literary hypertext tradition from Michael Joyce's *afternoon* [23] onwards, the story is hidden within a labyrinth of links and nodes. Mainstream Web sites of the 1990'ies are very different from this paradigm. They typically make any page easily accessible from any page with sets of navigation links. In such openaccess hypertexts, narration must be different. If any paragraph is easily available at any time, how can the discourse control the reader's knowledge any more? In a detective novel, we can always peek at the ending, but that is normally considered to be cheating. However, in an open hypertext, the outcome is an equally valid choice. Where the *Midway* film keeps the viewer wondering throughout the film whether Ballard will find the sunken *U.S.S. Yorktown*, the Web feature with the same title starts with the solution: "Robert Ballard (...) found the Yorktown on May 19, 1998."

This is not unique to the Web, however. *National Geographic Magazine* articles routinely start out with stating the outcome. A story always starts with a two-page spread containing a picture that sums up the main theme of the story, with the title and a lead paragraph printed onto it. The lead paragraph sums up the whole story. In the Midway article, the opening picture was an underwater photo of *U.S.S. Yorktown*, effectively disclosing that Ballard's expedition was successful.

This similarity between the Magazine and the Web sites can be explained by the fact that both are designed to be browsed. Many readers of *National Geographic Magazine* report that they tend to flip through the pages, looking at the gorgeous photographs, reading a few captions, and then, maybe, starting to read an article.

Where the narrative films keep the outcome secret, magazines and webs create interest by doing the opposite: by stating what has happened. It may be due to the media's different means of communication. Television films can hook the viewer by dramatic pictures in the opening without revealing the outcome. With text, it is easier to create the drama by stating the result. Print and Web articles rely on the principle of news articles, where the first paragraph summarises the whole story. This "inverted pyramid" principle of writing creates interest by insisting on the news value of what is to come. Further on, pockets of mini-narrative keep the interest: smaller enigmas, focalisations, and episodes. But also by the "how" enigma: If the outcome is disclosed: "The king died", one's automatic reply is "how?" And the answer to that can be infinitely detailed.

4.2.2 Collapsible narratives

The structuralistic strand of narrative theory from Propp [34] onwards focused their attention on the essential functions of narrative; turning points in the story, enabling some courses of action and excluding others. A part of this tradition is the search for minimal or universal narratives, enabled by the obvious fact that stories can be summarised. And when we summarise, we reduce stories to the functions of causality, causes and consequences, leaving out details that are not necessary for the understanding of the story line. When further summarising, we must leave out entire intermediate causal chains of events. It is exactly this aspect of narrative Genette is addressing, when he is able to reduce Proust's massive 7-volume Recherché du temps perdu into one sentence: "Marcel becomes a writer" [18, p. 30]. At the other extreme, any narrative can be fleshed out with more descriptive detail, as Barthes repeatedly points out [5, p. 108]; [6, p. 146]; [7, p. 82]. There probably is a limit to how short a narrative can be. Where the limit is is disputed, but a few words will probably suffice. At the other extreme, any narrative can be infinitely long. Any narration must find its place between the two extremes, choosing which whole sequences of actions to include, and which to exclude, but also choosing the number of details, the number of parts in each chain. In the Magazine article on Midway, we learn that a mini-submarine used in the search broke four times before it found the sunken Yorktown. In the film, only the second of these breakdowns is mentioned. A second example: In the article, the sonar searches are briefly mentioned, in the Web site, this slow and careful process is described in detail. The choice of such details is an important aspect of narration. However, we may strongly suspect most readers would never read the page on sonar searches in the Midway web, if they ever saw it. In an open-access hypertext, it depends on the reader's choices how many actions or events in a story she reads. Still, the story is the same, due to narrative's collapsible nature.



Figure 6. Opaque linking in the Midway Web site.



Figure 7. "Case folder" in Asteroid Web site.

4.2. Opaque linking

The *Dispatches* section in *Midway* Web contains 27 letters from a journalist who was part of the team that searched for the downed ships. The reports were posted on the Web server as they arrived from the journalist, giving a daily update as the expedition progressed (Figure 6).

Now they are available in this archive. On every page in the section, all 27 letters are available from a list in the left-hand column of the dates the letters were posted. Apparently, this is total multicursality. However, there is little hope of guessing what is contained on a page before arriving upon it. The link to each dispatch is signified by the date of the destination page, arguably the least telling aspect of the dispatch. To know what happened on that day, the reader has to browse it to know what the page contains. The discovery of the sunken ship, the climax of the story, happened on 22nd day, and is not marked differently from the others. A narrative is hidden behind the 27 links, but the reader has to navigate in them blindfolded. As we will see, the section could have given the reader overview and control of the narrative, but that would have required better signification of the destination pages. Reading in sequence, from the first day through the 27th is the only way to read Dispatches and be able to predict what you will find on the next page.

5. MULTICURSALITY?

We start to see that the dream of hypertext as a new, multicursal form of writing hardly is realised in our Web sites. The *Volcano* and *Midway* Webs have chains of pages linked after each other with presentation links. Although navigation links are provided, the reader's ability to control the sequence is obscured by a lack of signs of what is at the other end of the links.

Still, the writing is less fluid than the writing in the magazine articles or the editing of the television films. The ten transitional devices listed earlier are not found in the Web sites. Few attempts are made to bridge the gaps between pages and sections.

Video editing gives many possibilities of fluent sequences. However, in the Web sites, video is only used as a bonus, an asset to attract readers to the text. Short video clips are placed in the openings or ends of the chains of pages, not really related to the written text.

What we have called classificational link structure is the only device we have found where the label nonlinear seems fitting. Interestingly enough, the Web site that makes most use of it, the *Asteroid* Web, puts it within a strictly linear discourse. The *Asteroid* Web is a game in three parts. The first part is an introduction where the user goes through a series of pages where she gets an assignment: to identify what hit the earth in four cases of impact. The middle section has a classificational link structure, the four cases are identical in design, and are all linked to a library of different kinds of impact. (fig. 7) Sequence makes no difference in this part. When the right answer is found in all four cases, the third section opens, a series of congratulatory pages culminating in a short video clip. Multicursality is kept within a linear framework.

6. BEGINNINGS

Whether linear or multicursal, a feature in a mass medium invite readers to spend some time with them. This makes the opening a crucial point. The beginning should catch the reader's interest and encourage her to engage in the discourse.

In our material, there are media-specific differences in openings. Magazine articles open with a two-page picture with the title and an introductory paragraph printed onto it. The picture captures the theme of the story, especially when read together with the paragraph. Anchoring the picture in time and place, the introductory paragraph gives a summary of the theme of the article as well, and in the two more narrative stories, this is done by giving away the outcome at the very beginning. Then, overleaf, the articles move into an episode in an in medias res kind of opening.

We have seen that the film keeps this secret to itself until the end. Instead, the films all open with a prelude that focuses on what Barthes called the symbolic area, the main contradictions to be mediated. Tightly edited, the preludes contains many short shots, focusing on dramatic documentary pictures that encapsulate the conflict and drama of the story (the dangers of meteorites, the force of volcanoes, the horrors of war), but not the outcome. After about a minute, the *National Geographic Television* vignette is inserted, and a new introduction follows. In this introduction, the main characters are introduced, as the camera (and we as spectators) arrives on the scene by air, by boat or on foot.

The Web features focus neither on setting or basic conflict, but on the navigation and segmenting of the webs. They all open a "splash screen", a poster (or cover)-like page with the feature's title, maybe an illustration, and only one link, to what we will call the first page. This first page warrants a promise of content to come, and the user is urged to explore. It does little do entice the reader into the story, it just states that it is there, and which parts it has.

7. ENDINGS

Does a hypertext end? Yes and no, the theorists answer. Not in the same way as classic narratives end, closing off every single thread. Maybe, still, as we may have had all our questions, assumptions and expectations answered, although we have not read all the pages, is the answer of J. Yellowlees Douglas and Terence Harpold after analysing Michael Joyce's Afternoon [15, 20]. In Hypertext, George P. Landow states that "Readers cannot only choose different points of ending, they can also continue to add to the text, to extend it, to make it more than it was when they began to read" (58). If any of us has abandoned a book halfway through, or commented upon someone else's work, we might ask what is new about hypertext in these respects. But it is certainly true that as long as a hypertext does not have a single line, it cannot have a single ending. One of our Web features, Asteroids, does have a single line, and an ending. The other two not. In Volcano Web, and in parts of Midway, we have good enough overview of the text to realise that we have read all the pages. This is still not establishing ending like the documentary films and articles in our sample do. They all round off by pointing to the symbolic area, the larger perspective, the never resolved conflicts: the horrors of war, Volcanoes being constructive as well as destructive, the danger of meteorite impact being minute but having enormous possible consequences.

8. CONCLUSION

This analysis of multicursality in popular Web non-fiction showed that they are a lot more linear than one would suspect from a reading of hypertext theory. They are in fact authored as linear texts, except for sections describing objects. At the same time, they are a lot less fluent texts than magazine articles or broadcast documentaries. In our sample, the Web sites are a bit like small libraries, collections of descriptive texts, hierarchically organised, badly cross-referenced. The rhetorical convergence, far from resulting in a melange of advanced techniques form writing or film, has instead given collections of different kinds of text, with a strong emphasis on navigational links.

The analysis of these texts is not finished. Relations between image and language in the different media is not yet addressed, a grave omission when we consider the pride the National Geographic Society has in its photographs. This paper has also omitted many narratological issues that could have been addressed, such as focalisation and the interplay of time-planes. What I hope is that the paper has demonstrated that analytical tools derived from semiotics and narratology can heighten our understanding of how Web sites function as texts.

ACKNOWLEDGEMENTS

During this work, I have benefited from discussions with Espen Aarseth, Arild Fetveit, Gunnar Liestøl, Bjørn-Helge Mevik, Elli Mylonas, and Tanja Storsul. The US-Norway Fulbright Foundation and the Norwegian Research Council supported this research.

WORKS CITED

 Asteroids: Deadly Impact. National Geographic Society, 1997.http://www.nationalgeographic.com/asteroids/ index.html. Accessed February 14, 2001.

- Aarseth, E.J., Cybertext: Perspectives on Ergodic Literature. Baltimore: Johns Hopkins University Press, 1997.
- Allen, T.B., "Ghosts and Survivors: Return to the Battle of Midway", National Geographic Magazine, Vol. 195, no. 4, (April, 1999), pp.80–103.
- Barthes, R., S/Z. Trans. Richard Miller (Trans. of S/Z. Paris: Seuil, 1970). Oxford: Blackwell, 1993.
- Barthes, R., The Sequence of Actions, in The Semiotic Challenge. U of California P: Berkeley. 1994, pp. 136-148.
- Barthes, R., Textual Analysis of a Tale by Edgar Allan Poe, in The Semiotic Challenge. U of California P: Berkeley. 1994, pp. 261-293.
- Barthes, R., Wrestling with the Angel: Textual Analysis of Genesis 32:23-33, in The Semiotic Challenge. U of California P: Berkeley. 1994, pp. 95-135.
- Berners-Lee, T., Information Management: A Proposal. World Wide Web Consortium, 1989.
- Bolter, J.D., Writing Space: The Computer, Hypertext, and the History of Writing. Hillsdale, New Jersey: Erlbaum, 1991.
- Bolter, J.D. and R. Grusin, Remediation: Understanding New Media. Cambridge, Massachusetts: MIT Press, 1999.
- Breger, B. and P. Schnall, The Battle for Midway. Television Documentary, National Geographic Television, 1998.
- 12. Bush, V., As We May Think. Atlantic Monthly, 1945(176): pp. 101-8.
- Chatman, S., Coming to Terms: The Rhetoric of Narrative in Fiction and Film. Ithaca: Cornell U P, 1990.
- Crews, F.C., The Random House Handbook. New York: Random House, 1980.
- Douglas, J.Y., 'How do I Stop This Thing?': Closure and Indeterminacy in Interactive Narratives, in Hyper/Text/ Theory, G.P. Landow, Editor. Johns Hopkins University Press: Baltimore, 1994. pp. 159–188.
- 16. Eco, U., A Theory of Semiotics. Bloomington: Indiana University Press, 1976.
- 17. Freese, J.H., Rhetoric, in Aristotle in 23 Volumes. Harvard University Press: Cambridge, 1926.
- Genette, G., Narrative Discourse. Trans. Jane E. Lewin (Trans. of Discour du récit. Figures III. Paris:Seuil, 1972). Ithaca: Cornell University Press, 1980.

- Halliday, M.A.K. and R. Hasan, Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective. Language Education, ed. F. Christie. Oxford: Oxford University Press, 1989.
- Harpold, T., Conclusions, in Hyper/Text/Theory, G.P. Landow, Editor. Johns Hopkins University Press: Baltimore, 1994. pp. 189–222.
- 21. Heasley, M., et al., Return to Midway. nationalgeographic.com, 1998. http:// www.nationalgeographic.com/midway/index.html. Accessed February 14, 2001.
- 22. Heasley, M., et al., FALLOUT: Eye on the Volcano. National Geographic Interactive, 1998. http:// www.nationalgeographic.com/volcanoes/index.html. Accessed January 14, 2001.
- 23. Joyce, M., afternoon, a story. Cambridge: Eastgate Systems, 1987.
- 24. Kress, G. and T.v. Leeuwen, Reading Images: The Grammar of Visual Design. New York: Routledge, 1996.
- 25. Landow, G.P., The Rhetoric of Hypermedia: Some Rules for Authors, in Hypermedia and Literary Studies, P. Delaney and G.P. Landow, Editors. MIT P: Cambridge, 1991. pp. 81-103.
- 26. Landow, G.P., Hypertext: The Convergence of Contemporary Critical Theory and Technology. Baltimore: Johns Hopkins University Press, 1992.
- 27. Landow, G.P., ed. Hyper/Text/Theory. . Johns Hopkins University Press: Baltimore, 1994.
- Landow, G.P. and P. Delany, Hypertext, Hypermedia and Literary Studies: The State of the Art, in Hypermedia and Literary Studies, P. Delaney and G.P. Landow, Editors. MIT P: Cambridge, 1991. pp. 3-50.
- 29. Liestøl, G., Wittgenstein, Genette, and the Reader's Narrative in Hypertext, in Hyper/Text/Theory, G.P. Landow,

Editor. Johns Hopkins University Press: Baltimore, 1994.

- Manovich, L., The Language of New Media. Cambridge: MIT P, 2001.
- McEvoy, M., Violent Volcano. National Geographic Television, 1998.
- Nelson, T.H. A File Structure for the Complex, the Changing, and the Indeterminate. in 20th National Conference Proceedings. 1965. New York: Association for Computing Machinery.
- Nelson, T.H., Computer Lib/Dream Machines. Revised Microsoft ed. Redmond: Tempus, 1987.
- Propp, V.J., Morphology of the Folktale. Trans. Laurence Scott (Trans. of Morfologija skazki. Leningrad: 1928). Bloomington: Indiana University Research Center in Anthropology, 1958.
- Slatin, J., Reading Hypertext: Order and Coherence in a New Medium, in Hypermedia and Literary Studies, P. Delaney and G.P. Landow, Editors. MIT P: Cambridge, 1991. pp. 153–169.
- Weaver, K.F., Meteorites--Invaders from Space, National Geographic Magazine, Vol. 190, no. 3, (September, 1986), pp.390-418.
- Weinreich, E., Asteroids: Deadly Impact. Televison program, National Geographic Television, 1997.
- Williams, A.R., Under the Volcano: Montserrat, National Geographic Magazine, Vol. 192, no. 1, (July, 1997), pp.58–75.
- Yankelovich, N., N. Meyrowitz, and A.v. Dam, Reading and Writing the Electronic Book, in Hypermedia and Literary Studies, P. Delaney and G.P. Landow, Editors. MIT P: Cambridge, 1991. pp. 53–79.

Ramesh Srinivasan MIT Media Laboratory 20 Ames St., Room E15 – 020d Cambridge, MA 02139, USA +1 – 617 – 253 – 5108

ramesh@media.mit.edu

PhotoGlas

Roderick Ferguson MIT Media Laboratory 20 Ames St., Room E15 – 020d Cambridge, MA 02139, USA +1 – 617 – 253 – 5108

roderick@mit.edu

Befekadu Ayenew MIT Media Laboratory 20 Ames St., Room E15 – 020d Cambridge, MA 02139, USA +1 – 617 – 253 – 5108

befekadu@mit.edu

ABSTRACT

In this paper we describe *PhotoGlas*, a visual browser for news. *PhotoGlas* uses a web crawler to retrieve photographs from a number of pre-determined international-news web sites. The system then categorizes the images it retrieves based on keywords extracted from the article in which they are found. The images are presented as a montage that reflects a news topic. We describe the process by which categorization occurs and explore the potential of this system to present meaningful juxtapositions of news events to users.

Keywords

Montage, juxtaposition, visual browser, semiology, visual model building, photograph, news, categorization

1. INTRODUCTION

Writing in 1964, Barthes declared that "semiology aims to take in any system of signs, whatever their substance and limits; photographs, gestures, musical sounds, objects, and the complex associations of all of these, which form the content of ritual, convention or public entertainment: these constitute, if not *languages*, at least systems of signification."¹ For Barthes, language was cross-functional. In other words, it is determined by a set of complex connections between objects and other forms of expression, and that correspondingly meaning could be decoded by a system of connections that respond and interact with forms that are independent entities themselves.

This idea is confirmed by semiology, which states that codification is the formation of organized connections or relations with other forms by which matter or energy are organized. Semiology, then, is a relational process of codification by which networks of relationships develop to transform energy into concrete information. While the domain of semiology is large, we focus on visual symbols. First, we introduce Lindley's hierarchy

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

of codification for visual phenomena²:

perceptual level: the level at which visual phenomena become perceptually meaningful;

diegetic level: the means by which these phenomena are represented in two-dimensional space;

cinematic level: image compositions and their placement into an expressive artifact space;

connotative level: the level of metaphorical, analogical and associative meanings objects may have;

subtextual level: the specialized meanings of signifiers.

We believe that a key to unlocking the semiology of the Web is the development of systems that begin to represent symbols on both a diegetic and cinematic level based on the connotative and subtextual criteria of the information context around the photograph. This is particularly important because the diegetic role of an individual photograph as an expressive form is maintained along with its enhancement as a bridge to the information and concepts behind it (connotative). This can catalyze the development of systems that can start to represent the semantics, personalities, and meanings behind virtual events.

With the expansion of information available on the Web, there has not been a corresponding development of a structure by which the common ideas and topics among them can be effectively integrated. Specifically, in the area of news, while the presence of grassroots and international media has escalated on the Web, a thorough browsing of any single topic remains relatively difficult. There is no single place where one can find an exhaustive representation of all the perspectives, side-stories, and information related to a particular news topic. For example, while the Palestinian perspective toward recent confrontations in the Middle East can be found at the Electronic Intifada web site³, a contrasting Israeli perspective can be found at the Israeli Defense Force's web site⁴. We hypothesize, however, that a juxtaposition of the two articles could serve as a symbol that tells a more complete story.

The belief that such juxtapositions are useful is based on the arguments made by theorists such as Manovich, in his "Languages of New Media." Manovich argues that such uses of symbolism will drive the media of the future. He describes this movement as compositing:

The computer era introduces a different paradigm: This paradigm is concerned not only with time but space. . .Thus, in order to qualify as an example of montage, a new media object should fulfill two conditions: the juxtapositions of elements should follow a particular system; and these juxtapositions should play a key role in how the work establishes its meaning. . .The borders between different worlds do not have to be erased, the individual layers can retain their separate identity rather than being merged into a single space; the different worlds can clash semantically rather than form a single universe.⁵

Manovich uses the terms juxtaposition and montage. We invoke these forms as appropriate symbols for representing news on the web. Our definitions of these terms are as follows:

juxtaposition: positioning objects close together;

montage: an image made by the juxtaposition of photographs.⁶

We refer to juxtaposition as a description of form, and refer to montage as the process of combining photographs. We believe that the application of montage to the photographs found on such disparate news web sites as the Electronic Intifada and the Israeli Defense Force can serve as an engaging and thought-provoking introduction to complex events in the news. Jean Luc Godard, the new-wave filmmaker, referred to the world as a "vague and complicated system that we are all continually entering and watching."⁷ Montage would integrate space, through its juxtaposition of photographs that were previously separated from one another, and time, by representing these juxtapositions dynamically, as the photographs that constitute them change with the flow of news.

PhotoGlas, a web browser for news photographs, is our first attempt at building such a system. Our goal in this project is to categorize photographs retrieved in real-time from a multitude of news-related web sites and in turn present these photographs in meaningful juxtapositions for users.

2. VISUAL MODEL BUILDING

An important subgoal of the *PhotoGlas* project is to create a tool for encouraging critical thinking and reflection among its users. We believe this can happen through the process of visual model building. (Smith defines visual model building as the use of photography to construct a theory related to the information it conveys.⁸)

Smith's *Animal Landlord* project tests the ability of nature films to engage students in the development of reflective models that qualitatively describe the behaviors they observe. The results were contrasted with text-based (textbook) traditional learning methods. The experimenter concludes:

Visual events are rich with opportunities for students to pose questions and reflect on behaviors and processes. . As a result, imagery establishes a context for problem solving, for generalizing explanations from pictorial evidence 8

Smith describes the visual model building process in four major steps: (1) decomposition: identification of actors and in general central personalities/subcomponents of the presented photograph; (2) comparison: looking for similarities and differences between any given photograph and the others present in the film (or in our own conceptions/visual models); (3) linkage conjecture: analyses (based on previous knowledge and the appearance of the photographs) of the factors behind the choice and relative distances of the photographs displayed; and (4) model construction: generalization of causal models that explain the numerous conditions (time, setting, relationship between actors, etc.) that have made the photographs a reality.

Visual model building has motivated the design of *PhotoGlas*. We hypothesize that providing users with a juxtaposition of topically-related photographs will engage them in processes similar to those described by Smith. Through montage, users make comparisons between the setting, actors, emotions, and actions. These comparisons are used to discover links and common elements. It is from this perspective that Smith's final step, model construction, can occur. The commonalities and discrepancies that a user has perceived among the photographs lead the user to construct a story that captures the entirety of a news event.

3. PRIOR WORK

There have been other significant efforts in the use of montage for interactive storytelling that inform this effort. Murtaugh's ConTour system⁹ interactively suggests movies and pictures based on a user's choices of how the story should proceed. ConTour provides the user with the power to make these decisions by suggesting certain keywords based on the user's previous choices. These keywords are in turn linked to the movies and pictures from the system's database. ConTour is able to categorize its media objects with keywords through an analysis of what the context behind these objects, while still empowering the user to determine the story received. Tsarkova's North End *Chronicles* project¹⁰ (based on Boston's North End neighborhood) generates a number of potential story playouts using movie clips, then lets the user decide between three playouts that the system has decided are best for the story the user has experienced thus far. These playouts are juxtaposed with the movie clip currently playing to inform the user of the related branches his or her experience could follow. Together, this montage presents a picture of the user's immediate story environment.

4. CATEGORIZATION

We have constructed a simple web crawler to periodically check a number of pre-determined international-news web sites. Our crawler integrates, for example, the photographs it retrieves from a range of web sites that include The People's Daily News from Beijing, China, The BBC News, The Moscow Daily News, among others. We believe that combining these international news sources provides PhotoGlas with a more inclusive base from which it can depict news. We decided to build our own web crawler rather than use one publicly available because we wanted to optimize it for the domain of news websites, which often exhibit common patterns such as the format by which they present information. We constructed our crawler to account for the differences and intricacies between specific sites as well. For example, the crawler can recognize a photograph and its caption as existing in separate tables on the BBC site, while expecting them to be placed together on The People's Daily News site.

The web crawler is programmed to capture the associated caption and headline of an article in which a photograph is contained, along with the photograph itself. It is from this text that *PhotoGlas* is able to extract keywords that are used to categorize a given photograph.

We make two refinements to the keyword list before initiating the categorization process. First, we have implemented a stoplist. (Stoplists contain words that are structural and inherently provide little meaning to a given text.) We are using a predefined stoplist that was statistically generated from a *Time Magazine* corpus to filter our keywords¹¹. We have taken this abbreviated set of keywords and applied a simple Porter stemmer¹¹ to convert words with the same morphological form into the same "stem." For example, "aviator," "aviating," and "aviation" can be converted into "aviat."¹² These two steps are illustrated in Figure 1.



Figure 1. Flowchart detailing pre-categorization requirements

Having isolated relevant information associated with a given photograph, we then classify each photograph into a specific topic. The topics we have chosen are listed in Figure 2.

Topics		
War		
Terrorism		
Poverty		
Medicine		
Business/Industry		
Justice/Police		
Entertainment/Media		
Nationalism/Patriotism		
Technology		
Women/Feminism		
Religion		
Politics		
Government/Leadership		
Discrimination/Inequality		
Nature/Environment		

Figure 2. Topics chosen for categorization

To do the actual categorization, we pre-assembled a list of keywords associated with each topic. This was done using WordNet®¹³, an on-line lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory. English nouns, verbs, adjectives and adverbs are organized into synonym sets, each representing one underlying lexical concept. For each of our topics we have assembled a set of WordNet-derived synonym and descriptive keywords.

To make the assignments between photograph text-derived keywords and topic-derived keywords, we implemented a simple matching algorithm.

For each photograph *i* in the database, let K_i be the set of keywords associated with the photograph. For each topic *j*, let T_j be the list of keywords associated with the topic. *PhotoGlas* calculates the correspondence between each photograph and topic using the ratio, R_{ij} .

 $R_{ij} = (K_i \cap T_j)/T_j$, where $K_i \cap T_j$ refers to matching keywords found in the intersection of the two sets.

The correlation between a photograph and each of the 16 topics is expressed through an array of these ratios. We sort these arrays based on descending ratio values. The greatest ratio determines the topic to which a photograph is assigned. In the case of a tie, we assign the photograph to multiple topics.

We were interested in expressing relationships between topics as well. With this motivation, we asked a number of prospective users to try to express which topics listed in Figure 2 were most similar to each other. Our matrix of inter-topic similarities, detailed in Figure 3, is a result of this study.

Торіс	Related Topic
War	Nationalism, Religion, Terrorism
Terrorism	Nationalism, Religion, Justice/Police, War
Poverty	Discrimination/Inequality, Nature/Environment, Medicine
Medicine	Technology, Business/Industry, Nature/Environment
Business/Industry	Technology, Justice/Police, Medicine, Nature/Environment
Justice/Police	Business/Industry, Discrimination/Inequality, Politics, Entertainment/Media
Entertainment/Media	Business/Industry, Technology, Politics
Nationalism/Patriotism	Government/Leadership, Politics, Discrimination/Inequality, War
Technology	Business/Industry, Medicine, Entertainment/Media
Women/Feminism	Discrimination/Inequality, Nationalism/Patriotism, Medicine
Religion	Nationalism/Patriotism, War, Politics, Terrorism
Politics	Government/Leadership, Justice/Police, Discrimination/Inequality, War
Government/Leadership	Nationalism/Patriotism, Justice/Police, Politics, Discrimination/Inequality

Discrimination/Inequality	Justice/Police, Politics, Government/Leadership, Women/Feminism, Poverty
Nature/Environment	Business/Industry, War, Technology, Politics

Figure 3. Inter-topic similarity ontology

5. INTERFACE

The *PhotoGlas* interface presents four topically-related spheres, a central one orbited by three satellites. Each sphere contains an identifying label for its topic. Photographs within each topic cluster are presented as a projection across the surface of a sphere. The shape of a sphere was chosen for its globe-like form. (We hope to remind users that when they are browsing in *PhotoGlas*, they are interacting with the world globally.) Selecting an individual photograph on any of the spheres takes the user to the web page from which that photograph was extracted. An example of the interface is displayed in Figure 4.



Figure 4. A screenshot of the PhotoGlas interface

The central sphere reflects the focus topic chosen by the user while the three spheres in the periphery reflect topics the system has identified as related to the focus. This identification is based on a hand-crafted ontology of topic similarities that we have created. The interface does constrain this ontology, as it allows a maximum of three satellite spheres to be displayed at once. We are investigating alternate interface forms that can better capture inter-topic relationships.

Selecting a satellite sphere causes that sphere to move to the center and for it to be surrounded by new satellites. This

provides a means for the users to navigate the topic space and refine their browsing interests. For example, if the topic "war" shows photographs that do not quite reflect a user's interests but is still related to what he or she has conceptualized, the peripheral topics of "nationalism" or "terrorism" might suffice.

6. IMPROVEMENTS

What we have accomplished thus far only reflects our initial efforts to build an effective visual browser. There are a number of ways in which we plan to improve *PhotoGlas*:

- The assignment of photographs to "relevant" topics is fundamental to *PhotoGlas*. It is important that the topic areas we choose are indeed what users find most relevant, and that the photographs assigned are reflections of the topic. There has been significant research done on news ontologies and what users globally find newsworthy. We hope to define topics based on these results rather than just our intuitions.
- The decision to assign a photograph to a single topic merely because its ratio was highest may be shortsighted. We suspect that a photograph can potentially evoke a number of different topics and perspectives. We will to investigate the idea of linking photographs retrieved not only to multiple categories but also perhaps directly to each other. This could be achieved by using matches between the keywords themselves.
- Instead of creating groupings solely based around topics, groups can also derive from a specific keyword from a particular photograph. The choice of keyword could then generate another group of photographs, creating a juxtaposition of photographs that can represent a particular news keyword.
- We would like to develop a more rigorous means of connecting the components of our topical ontology. We plan to design an experiment to reveal which topics are indeed "closer" to others.
- Finally, we intend to investigate different interface forms that may provide a better sense of the context of a photograph. It may be that a static juxtaposition may not be an ideal form for the expression of context.

7. EVALUATION

We have articulated our goal of constructing a tool to present users with meaningful juxtapositions of news photographs. We also stated a subgoal was to encourage critical thinking and reflection among users.

In *PhotoGlas*, we have created a technology platform that allows us to experiment news montage. The question that we hope to focus on is which juxtapositions are actually meaningful to users? This project allows us to begin our search for the answer to this question.

In the spirit of Smith's Animal Landlord project, we believe that *PhotoGlas* can serve as an appropriate environment from which the four steps of decomposition, comparison, linkage conjecture,

and model construction can be taken. We intend to examine whether or not (and if so how) users go through the visual model-building process when interacting with *PhotoGlas*. Specifically, we hope to investigate whether or not Smith's theory of learning through reflecting can apply to a situation where there is no explicit task at hand.

We are also interested in determining what factors make a photograph evocative to its viewers, particularly when it is placed in a montage. We plan to address the problem of how to train *PhotoGlas* to select photographs that users find powerful.

Finally, *PhotoGlas* may benefit from incorporating advancements in the field of photograph recognition. The ability to capture certain descriptive aspects of a picture (such as the identity of people or a physical location) or aspects of its composition may further *PhotoGlas's* goal of providing users with a more powerful combinations of photographs. This raises the possibility of enabling computers to associate a topic with a photograph without basing this association solely on the text accompanying the photograph.

8. ACKNOWLEDGMENTS

We would like to thank Walter Bender, for his guidance and active assistance on this project. This work is sponsored by the MIT Media Lab's Information: Organized Research Consortium and *eircom*.

9. REFERENCES

 Barthes, Roland, *Elements of Semiology* (trans. Annette Lavers & Colin Smith). London: Jonathan Cape ([1964] 1967).

- [2] Lindley, Craig, "A Computational Semiotic Framework for Interactive Cinematic Virtual Worlds": <u>http://wheelie.tees.ac.uk/users/p.c.fencott/</u>
- [3] Electronic Intifada site: http://electronicintifada.net/new.html
- [4] Israeli Defense Fund site: http://www.idf.il/english/news/main.stm
- [5] Manovich, Lev, Languages of New Media, MIT Press, March 2001, p.147
- [6] Word Dict site: http://www.dict.org
- [7] Godard, Jean-Luc, Son + Image, edited by Raymond Bellour (New York: Museum of Modern Art, 1992) p. 171
- [8] Smith, Brian K. & Blankinship, Erik, *Imagery as data:* Structures for visual model building In Proceedings of Computer Support for Collaborative Learning 99, p.549-557
- [9] Murtaugh, Michael L, *The Automatist Storytelling System* (MIT Media Lab, 1994)
- [10] Tsarkova, Natalia, Dynamic Collage: An Interface for Digital Stories (MIT Media Lab, 1997)
- [11] Marlow, Cameron A, A Language-Based Approach to Categorical Analysis (MIT Media Lab, 2001)
- [12] Porter, M.F, An algorithm for suffix stripping. *Program* 14(3), (1980), p.130-137
- [13] WordNet site: http://www.cogsci.princeton.edu/~wn/

Margarete Jahrmann / Max Moswitzer http://www.konsum.net Konsum Art-Server Vienna, linx3d@konsum.net

===== KONSUM Gamearchitecture ==== conceptual arts programms based on network processes ====== _____

Konsum Art-Server nTRACKER

http://www.konsum.net/ntracker

- machinima movies
- data-Objectiles
- data transform-software

preambel>- - reflecting the structures of the net> data and protocollinherent functions of the networks are processing and permantly restructuring databodies. A pseudo-netcredibility is set up by the TRACES you continously leave on the net. Tele-workers have to proofe their online-time and are paid by logfiles. Neo-Keynsianisme is no more necessary- LOG-Functions are the basis of the net-protocolls. Lets start to make it obsolete-and smile on it.



Zagreb 26thYouth Salon 2001



<*c*'est mon choix!>: DJing lecture events combine and demonstrate in a showcase the use and live combination of Konsum/Server-Tools. In a lecture situation the "mediator superfem" is presenting her autotheory (based on the simultanously projected work). The "mutator moswitzer" is operating the nTracker soundgeneration and navigating live through 3D multiuser-objectiles. Online users are triggering the in situ mix with screen-projections of machinima films. Participant, it is not your free choice, you are tracked in our Information Technologies Society: http://www.konsum.net/2000/live.html





2. All the TOOLs provided on the ntracker and other pages on konsum can be seen as SITUATIONIST ironic TOURNeMENT instead of deathmatch Tournament (digital waffen-lager) to curve out alternative fake bodies by code-text. We were using GAME-Engines for that and creating 3D playable Multiuserenvironments based on databodies. *see linx3D. DATA-OBJECTILES can even be printed as 3D Real-Life objects and display your online live of f.ex 1month. http://www.konsum.net/ntracker/coop

1. In HEDONIST art situations- called DJing lectures or live Traceroute-performancesas described on our OWN independent server KONSUM.NET we can start a pleasurable reflection and relativation of all these circumstances of online live and living from e-mail. http://www.konsum.net/server/traceroute2000





3.MACHINIMA - FILMs Btw: by our software even a film of your databody can be rendered in realtime- ordinary it is displayed on the net and can be transfered very quickly - because it is of small data-size and everyone who wants to see it can install it (attention: grafix engine required first) on her computer. - but you also can play a copy on a video-casette. SEE our demo-machinima-movies of data-bodies. http://www.konsum.net/areal

Margarete Jahrmann / Max Moswitzer http://www.konsum.net Konsum Art-Server Vienna, linx3d@konsum.net



DATA Transfer/Objects_ MACHINIMA GAMING Events

KONSUM Gamearchitecture

nTracker is no Web-Art. It is a complex connex-ware integrating tools/and software modules located on different Konsum-sites with data-filters transforming data into aesthetic experiences into an Event-matrix.

nTracker uses everything that goes through your network card (tcp-dump- your datatrash on the machine) and processes the data for hedonist art SituationisMe.

It is inspired by the Demo-scene: each gameCrack has also an tracker for creating the sounds following a new technological concept of minimodules and time triggers. This is helpful for traceroute timestamps.We use it that way for live LAN-parties.

A MACHINIMA is also always displayed at the ntracker events. This is a realtime 3D game-engine based movie rendered live. So we can integrate live "Frequences" of the renderer Hardware as well into the visuals.



Die Daten und protokollinhärenten Logfunktionen des Internets bilden die "Proofable Units", beweisbare Einheiten, die zur Wertsteigerung einer Location im Netz beitragen können.

Eine Wertstandardisiserung im Netz könnte durch Datentransparenz erreicht werden. All diese Prozesse aus dem Internet können mit serverseitigen Skripts visualisiert werden. Server Protokoll Experimente und ihre Visualisierung. Der Transfer von symbolischem ins ökonomische Kapital erfolgt durch Datavatars. Die Transparenz der Datentransfers und Serveraktivitäten ist eine Voraussetzung um als psychoorganische Systemkonfiguration/Mensch oder grössere Dateneinheit "Netcredibility" nachweisen zu können.



Diese Thematik hatten wir bereits mit unserer Multiuser/Traces Arbeit linx3d aufgegriffen.

Nun geht es aber darum diese proofable units zu transferieren- in lustvoll erlebbare Umgebungen zu bringen.nTRACKER RELEASE ist ein- sound Transfer zum lustvollen Erleben von Data-TRACKs in Live/Situationen. Ausgangspunkt sind alle Daten, die über die Netzwerkkarte eines Computers kommen. Das Protokoll, das dieses ohnedies festhält heisst tcp-dump.

KONSUM Gamearchitecture

Ein Ausschnitt aus diesem Datenwulst – den Netzwerkdenken heisst Ausschnittdenken-wird in der System eigenen künstlerischen Logik zum Transfer in ein Dataobjekt genutzt. Zuvor hatten wir soundbasierte Transformationen und Traceroute für DJing lectures verwendet.



Margarete Jahrmann und Max Moswitzer http://www.konsum.net Konsum Art-Server Vienna, linx3d@konsum.net



Die Data-objekte können nun als frei flottierende Objektile nach Deleueze Guattarai im Datenraum wirksam werden. Sie werden nicht als Datavisualistation verstanden sondern als ästhetische Transformation von Daten, die auch als SPIEL-Architektur verwendet werden kann. C'est pas mon choix- sie haben keine Wahl, wie ihre Multiuserumgebung aussieht- sie wird aus Datenbeständen prozessiert. Das Lesen des 3D-Datenobjektes als Informationsumgebung ist nicht relevant. Es ist einfach ein Faktum den eigenen Logfile- DUMP – also Müll als eine Multiuserumgebung anzubieten. Der eigenen Datnkörper wird nicht geschützt sondern allen als Spiel-Umgebung zur Verfügung gestellt.

All jene die nicht spielen wollen können sich auf KONSUM- auch ein Machinima Movie des DatenTransform Objektes reinziehen. Sie können im gemütlichen machinima-Movie einen live-Film Mitschnitt von Aktionen anderer user in diesem Datenraum ansehen- ähnlich

wie die KIDS und Gambler Gemeinde sonst anderen ihre besten Shooter-Aktionen demonstriert. Der Machinima Raum öffnet noch eine weitere Dimension der Online- Echtzeit movie Generierung im absoluten Low-Budget Kontext und kommt ausserdem aus einer subversiven Tradition der Demo- und CRACK Szene, die einfach ihre gecrackten Softwarespiele in Intro-filmen, die ähnlich einem Graffiti Schriftzug zur verschleierten Identität funktionierten-kurz MACHINIMA ist ein schöner polyvalenter Assoziationskontext. KONSUM Gamearchitecture

Für eine Installation wird neben der Projektion des Live- Multiuserspieles und eines MACHINIMA movies auch ein 3D Objekt der Daten-Transfom Objekte belichtet werden. Es kann als Erinnerungsstück, als haptisches Gefühls und Genusselement dienen, das sich vom technologischen Interface emanzipiert hat.

Das "Ertasten" einer Momentaufnahme eines TCP-DUMP Datentransformobjektes könnte als Erlebnismehrwelt adventureride-maessig erlebt werden.

(no e-cash but fun):

ironisches Tournament und hedonistische Djing Lectures. Solche Live-Performances (Sound-Visual-Text) mit Datenströmen sind seit 1996, seit wir den KONSUM-ArtServer für experimentelle Datentransfers gegründet haben, unser FORMAT auf Medien- Kunstfestivals (z.B. Rotterdam DEAF Festival 1997, Osnabrück EMAF 1998, ARCO Electronico Madrid 1998, Ars Electronica 1999, steirischer herbst 1999, net_condition ZKM Karlsruhe 2000, World-information.org Bruessel 2000, Zagreb Youth Salon 2001.) see http://www.konsum.net





Bisher haben wir stets versucht diverse CRACKS im ästehtischen und kulturellen Feld von Game Engines zu realisieren, wie etwa "linx3D" das für das ZKM Karlsruhe, ICC Tokio und Mecad Barcelona, das als Installation in den Museen mehrere Ausstellungsorte verbindet und vom Internet aus benutzbar ist. http://linx3d.konsum.net

Das Game wurde in eigens gebauten "Konsolen" inszeniert. Vgl. dazu (siehe auch Cover) des Katalog-Buches "netCondition" erschienen bei MIT Press 2001, Hg. Timothys Druckrey, Peter Weibel).

Unsere "Produkte" waren stets live-events, bei denen live generierte ONLINE Sounds und Spiel-Umgebungen in Echtzeit3D für Visuals benutzt werden- z.B. eine LAN-Party wie bei Synworld im Messepalast Wien 1998. mehr dazu auf http://www.konsum.net/areal
Cypher: Cyber Photographer in Wonder Space

Takahiro Nakai[†] Makoto Tadenuma[†] [†] ATR Media Integration & Communications Research Laboratories 2-2-2 Hikaridai Seika-cho Soraku-gun Kyoto 619-0288

Japan

Shoji Tanaka^{††} Keiko Nakao[†] ^{††}presently: Mitsubishi Electric Corporation Information Technology R&D Center 5-1-1 Ofuna Kamakura-shi Kanagawa 247-8501

Japan

(formerly: ATR Media Integration & Communications Research Laboratories)

+81-774-95-1465 +81-774-95-1460

+81-774-95-1472

tnakai@mic.atr.co.jp tadenuma@mic.atr.co.jp gon@isl.melco.co.jp nakao@mic.atr.co.jp

+81-467-41-2098

ABSTRACT

In order to make it possible for us to easily create a virtual world simply by using our imagination with building blocks and to take photographs of ourselves traveling in the wonder space with a beautiful composition, we have developed a system called "Cypher (Cyber Photographer in Wonder Space)". With this system, users can blend the real world with a virtual world by using their aesthetic sense. This system is expected to be applied not only to the amusement field but also to support tools for city planning or layout simulations involving furniture and so on.

Keywords

building blocks, fusion system, virtual world, 3D objects, CG, VRML, pose recognition, P type Fourier descriptors, circumference curve, physical image, compositional knowledge

1. INTRODUCTION

Is it possible to create a virtual world easily without special technical knowledge simply by using our imagination with wooden blocks (i.e., the same building blocks that children play with) and to take photographs of ourselves traveling in the wonder space with a beautiful composition? In order to realize such a dream, we have developed a system called "Cypher (Cyber Photographer in Wonder Space)". Cypher is a fusion system that consists of virtual worlds and photographs of the real world.

2. OUTLINE OF THE SYSTEM

Figure 1 shows the appearance of the Cypher system and Figure 2 shows the outline of the system. This system has four kinds of virtual worlds as shown in Figure 3. They include 1) the Japanese world in the Edo period, 2) a rabbit world, 3) a cosmic world, and 4) a haunted world.

First published at COSIGN-2001'

10-12 September 2001, CWI,

The Netherlands.

When using the system, a user first chooses a favorite world, and then arranges crystal blocks on an exclusive table. Each crystal block contains a model of a building or tree, etc., and each model is assigned to a virtual object. There are eight crystal blocks in total. They include one human figure (assigned to the user himself/herself), two vehicles, three buildings, and two trees. Multiple users can participate in this system at the same time, and try their own arrangements. The situation of each arrangement on the table is reflected on the screen in front of the table in real time. These virtual objects are arranged in the virtual world according to the positions and the directions of the corresponding crystal blocks. The user can proceed with manipulating the arrangement while confirming the quality of the arrangement. After the virtual world is constructed, the user then places a human figure block where he/she wants to appear in the virtual world and instructs the system to take a picture of him/her. Then, the system extracts the image of the user from the picture, and analyzes the physical pose of the user. This system maintains a database of about 200 painting masterpieces and also the physical poses of the main figures contained in these masterpieces. From this database, the system selects the masterpiece containing the nearest physical pose to the user's, and synthesizes a photograph by placing the user's image at the position of the human figure block in the virtual world based on



Figure 1. Appearance of Cypher



Figure 2. Outline of Cypher



The Japanese world in the Edo period

A rabbit world

A cosmic world

A haunted world

Figure 3. The Four Kinds of Virtual Worlds in Cypher

the compositional knowledge of the selected masterpiece. The resulting photo is printed out and presented to the user.

3. CONSTRUCTION OF THE VIRTUAL WORLD

As is shown in Figure 1, this system has an exclusive table of about 95cm in height at the center, and the top surface of the table is made of a transparent square glass plate about 110cm on one side. The eight crystal blocks, each about 10cm on one side, can be arranged on this table. In the following, how a virtual world can be constructed is explained.

3.1 Measurement of the 3D Position of Each Block

Figure 4 shows how the 3D position of each block is measured. Each block has an ultrasonic oscillator at the center of the top surface, and each oscillator has a different ID code. Four ultrasonic sensors are installed on the ceiling at the position of each apex of the square room; the diagonal is about 130cm. From these sensors, eight different ID codes are sent sequentially by infrared rays. Then, the ultrasonic oscillator that has the same ID code sends back an ultrasonic pulse. Each ultrasonic sensor receives the pulse, and knows the distance from the oscillator by the corresponding delay time. The 3D position of the ultrasonic oscillator is determined based on the principal of three-point measurements. Therefore, the 3D position of each ultrasonic oscillator is determined. In these measurements, it is possible to determine all 3D positions only with three ultrasonic sensors. In this case, however, there is the possibility that one direction might be cut off since all users place only the blocks in their hands. This is why we use four ultrasonic sensors. When an ultrasonic pulse reaches all four sensors, we can select the three sensors that make the error minimum, so that the precision of the measurement can become higher.



Figure 4. Measurement Method of a Block's Position

3.2 Measurement of the Direction of Each Block

As shown in Figure 5, a square fluorescent substance that shines palely when responding to an ultraviolet ray is stuck at the bottom of each block. On this square surface, a black square marker of about 1.3cm on one side is also stuck on the surface at a position that deviates about 2cm from the center of the square surface towards the front of the block. On the other hand, at the inside bottom of the exclusive table, a camera is installed that takes a picture of the current image on the upper glass as well as black lights that emit ultraviolet light. The picture taken by this camera contains eight white squares and black markers clearly as shown in Figure 5.

The system extracts each white square that contains a black marker from the picture taken by the camera. In this process, these extracted white squares' edges have many irregularities because of noise. Accordingly, we find the straight edge lines of the white squares using Hough transforms. Then, the four intersection points of the four straight lines are identified as the apexes of each extracted white square.

Next, the system performs pattern matching between each extracted square and the patterns rotated around the z-axis as shown in Figure 5. The rotation angle that gives the best matching determines the direction of the front of the block.

3.3 Generation of Objects

Based on the positions and the directions measured by the above methods, 3D objects corresponding to the individual blocks are generated by CG, and are arranged into the selected virtual world. In the generation of the CG, VRML is employed because of the ease it allows in CG development. Four different types of objects have been prepared according to the four virtual worlds. For example, a vehicle block corresponds to a horse or a large two-wheeled cart in the Edo period world, but it corresponds to a flying car in the cosmic world.



Figure 5. Measurement Method of a Block's Direction

Moreover, because the complete lack of movement in the constructed virtual world can create boredom, we give some objects some movements or we add some additional moving objects. For example, in the case of a vehicle object, we make it move around the center of the block, or in the case of a building object, we make a little girl walk around the building or make the door open and close. By these movements, even after the layout of all objects is fixed, many movements exist in the displayed virtual world, and we can look at it without losing interest.

4. PUTTING THE USER'S IMAGE INTO THE WONDER SPACE

With the method mentioned above, a virtual world can be constructed easily according to the user's imagination. But the interaction between the user and virtual world is slightly poor. Therefore, we add a function to this system to take a souvenir photo of the user traveling in the wonder space with a beautiful composition. With this function, a model of a beautiful composition is taken from a database of compositions of about 200 ancient and modern famous painting masterpieces. Using this database, the user can obtain a souvenir photo with a beautiful composition, even if he/she has no knowledge about the compositions of paintings or photos. In the following, this process is explained.

The basic system of this process is "The Cyber Photographer" [1], which has become a part of the system through name association. This is a technique that extracts the main objects from a picture panoramically taken with a beautiful composition. At this time, the main object is restricted to the user's figure. Therefore, we do not need to take a picture panoramically expressly, and can simplify the extraction process.

As mentioned above, after placing a human figure block where he/she wants to appear in the virtual world, first, the user takes a picture of himself/herself in front of a blue background. Then, the system extracts only the image of the user by a simple chroma key process.

Next, the system performs a pose recognition process about the physical image of the user. For this pose recognition, the figure of a circumference curve of the physical image is used. This closed circumference curve is described by P type Fourier descriptors [2].

The descriptors are obtained as follows. First, the closed circumference curve is divided into N segments. Next, the lines of each segment are normalized to vectors of length 1, and placed on a complex plane. Then, this sequence of complex numbers is Fourier transformed. These numbers in the frequency domain come to be the P type Fourier descriptors. By using these descriptors, we can describe the figure independent of the size, and make the position of the terminal of the reconstructed curve coincide with the position of the original curve. Moreover, we can reconstruct a smooth circumference curve while well preserving the rough figure by taking out the lower frequency components up to n (<N). Here, we use N=512, n=40.

On the other hand, for the circumference curves of the main figures of the approximately 200 masterpieces, P type Fourier descriptors of N=512, n=40 are stored as a database. By calculating the summation of the square errors between these P type Fourier descriptors, the figure in the database that gives the smallest summation value is selected as the pose nearest to the user's.

In the database, the area ratio of the human figure to the entire screen and the position of the center of gravity of the figure are also preserved. In order to make this information coincide, the size and the position of the user's image are converted. During the conversion, the camera is assumed to be on the normal line of the human figure block, in order to minimize the number of variable parameters of the camera. Accordingly, the conversion is carried out by the change of the distance to the virtual camera and horizontal pan and vertical tilt.

Figure 6 is an example of a souvenir photograph.

5. CONCLUSION

This system "Cypher" makes it possible for a user to create a virtual world simply through his/her imagination with toy building blocks, and to take a souvenir photograph of



Figure 6. Example of a Work

himself/herself traveling in the virtual world with a beautiful composition. With this system, the user can blend the real world with the virtual world using his/her aesthetic sense.

This system can be applied not only to the amusement field but also to support tools for layout simulations of rooms in a virtual house, or for layout simulations of buildings and so on in city planning, by replacing the background images and objects.

From now on, we plan to increase the kinds of contents and improve the quality of the contents. We also want to increase the picture quality of souvenir photographs.

6. REFERENCES

- Shoji Tanaka, Jeffery Chi Yin Ng, Keiko Nakao, and Yuichi Iwadate. The Cyber Photographer. Technical Report of IEICE, IE99-143 (February 2000), 41-45.
- [2] Yoshinori Uesaka. A New Fourier Descriptor Applicable to Open Curves. Transaction of IEICE Vol. J67-A, No. 3 (March 1984), 166-173.

The Link is the Data

On Realisations, Representations, and the Link Betwixt

Werner Kriechbaum Gerhard Stenzel IBM Deutschland Entwicklung GmbH Schönaicher Strasze 220 D-71032 Böblingen, Germany

{kriechba,stenzel}@de.ibm.com

ABSTRACT

The paper shows how links can be used to imprint structural and supplementary information from the representation of multimedia data on the realisation and thus enable a structure-based navigation. The authors suggest that similar linking mechanisms can and should be used to connect descriptive metadata with the essence.

1. INTRODUCTION

Most if not all multimedia data exist or can be expressed in two equivalent forms, either as a symbolic representation, or as a realisation. This duality is most obvious for audio material but holds for audio-visual material or "non-traditional multimedia" data, like e.g. buildings or proteins, as well. Examples of representations are the score of a piece of music, the storyboard of a movie, and the layout of a building - corresponding to the realisations of an audio recording, the filmed movie, and the building either built in reality or virtual reality. In many cases, the symbolic representation captures all but the "emotional" aspects of a realisation. The loss of this aspect is often compensated by the fact that the symbolic representation either contains additional structural information or makes it comparatively easy to derive this information. Therefore the representation is a powerful tool to ease the navigation and the interpretation of the corresponding realisation, given the relation of elements from the representation to elements of the realisation can be established and encoded.

2. REPRESENTATION AND REALISATION

Representation and realisation are different manifestations of the same content and can – at least in principle – be generated from the respective other form (Fig. 1): A representation can be rendered to create a realisation and a realisation transcribed to produce a representation. Many attempts to make multimedia data more accessible for indexing, search, or navigation try to derive a structure from audio or video

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands.



Figure 1: The relation between representation and realisation

streams (see e.g. [16]). Despite tremendous advances in rendering as well as in transcription technologies (for recent examples in the visual domain see [8],[10],[13],[17]) the automatic extraction of a symbolic representation from multimedia data is still much harder than the generation of a faithful realisation. And even if an accurate transcription can be achieved much of the underlying structure of the data is lost: State-of-the-art speech recognition systems still rely on their user to explicitly supply punctuation marks. But in cases where a symbolic representation is available, the structure can be extracted from the representation and imprinted on the realisation.

Consider for example the audio or video recording of the staging of a theatre play. The text of a play like Shake-speare's Hamlet has a rich structure built from acts, scenes, and dialogues (Fig. 2) [18] that can be made explicit by markup.

<ACT><TITLE>ACT I</TITLE> <SCENE><TITLE>SCENE I. Elsinore. A platform before the castle.</TITLE> <STAGEDIR>FRANCISCO at his post. Enter to him BERNARDO</STAGEDIR> <SPEECH>



Figure 2: The simplified TEI document type definition for a play

<SPEAKER>BERNARDO</SPEAKER> <LINE>Who's there?</LINE> </SPEECH> <SPEECH> <SPEAKER>FRANCISCO</SPEAKER> <LINE>Nay, answer me: stand, and unfold yourself. </LINE> </SPEECH> <SPEECH> <SPEAKER>BERNARDO</SPEAKER> <LINE>Long live the king!</LINE> </SPEECH> <SPEECH> <SPEAKER>FRANCISCO</SPEAKER> <LINE>Bernardo?</LINE> </SPEECH> <SPEECH> <SPEAKER>BERNARDO</SPEAKER> <LINE>He.</LINE> </SPEECH> <SPEECH> <SPEAKER>FRANCISCO</SPEAKER> <LINE>You come most carefully upon your hour. </LINE> </SPEECH> . . . <STAGEDIR>Enter HORATIO and MARCELLUS</STAGEDIR> <STAGEDIR>Exeunt</STAGEDIR> </SCENE>

```
<SCENE><TITLE>SCENE II. A room of state in the
castle.</TITLE>
<STAGEDIR>Enter KING CLAUDIUS, QUEEN GERTRUDE,
HAMLET, POLONIUS, LAERTES, VOLTIMAND, CORNELIUS,
Lords, and Attendants</STAGEDIR>
<SPEECH>
<SPEAKER>KING CLAUDIUS</SPEAKER>
<LINE>Though yet of Hamlet our dear brother's death
</LINE>
```

Tags like ACT or SCENE allow structure based navigation ("Go forward one scene") and structure aware queries ("Which speaker referred to Hamlet for the first time and in which scene?"). Imprinted on the audio recording the same navigation is possible for the realisation – quite a difference from the customary navigation based on track numbers and timecodes. In addition the text of a play contains a wealth of stage directions that provide additional information. The original stage directions of a play can never be derived from the realisation, even if some sophisticated analysis might be able to interfere them, it is not possible to ascertain that the performance honours the original stage direction without analysing the representation. A similar argument applies to the structuring of music. It is possible to structure an audio recording according to derived dynamical patterns, but without access to the score it is not possible to structure the realisation according to the dynamical patterns intended by the composer. Like structure stage directions can be used for navigation ("Go to where Horatio and Marcellus enter stage") and query but in addition some of the stage directions can provide helpful hints for automatic transcription tools: For a person tracker it is rather useful to know who is - or should be - on stage and who not.

Often a representation has more than one structure: A printed book has a surface structure (pages) as well as a deep structure built from parts, chapters, subchapters, footnotes and cross-references. Finding and addressing specific structures (pages, chapters, etc.) in its audio book realisation is either time consuming, complex, or impossible. Whereas this is only inconvenient for the casual user of an audio book, the capability to search, navigate, and cite conforming to the printed edition is essential for blind students when working with audio text books. Linking representation and realisation solves many of these problems: The structure information can be used to segment the audio signal in meaningful units like sentence, paragraph or chapter, and standard text-mining technologies can be used to locate sentences of interest. Even, if the transcript is unstructured in the sense that it does not contain formal markup, punctuation can serve as means to partition the document and thus create a very basic kind of structure.

3. COMPLEX STRUCTURES

In most cases the document structure of a play or a book seems to be rather simple: There is a surface structure that refers to the physical appearance of the printed representation (page and line numbers) and a structure that describes the content as such (acts, scenes, speeches or chapters, subchapters, paragraphs). More complex structural relationships exist but these are more apparent in music and will therefore be discussed in the context of classical European music. Music has a complex temporal organisation and a rich semantic structure that defines a natural segmentation for the audio stream. Like in images, in music many features are characteristic for segments, vary from segment to segment, and become meaningless when averaged over all segments. For most, if not all music, the structure is not arbitrary but conforms to one of a small set of possible pattern [11]. The sonata form, or better sonata forms [15] since it is a set of closely related patterns, is one of the most important patterns of Western music in the Classical Era and a highly simplified sketch of the structure of this form is depicted in Fig. 3. The sonata form is built from at least three pieces: An optional Introduction, an Exposition, a middle part (called "Durchführung" or Development), a Repeat and an optional Coda. The Exposition is not an atomic structural element but built from the succession of a First Theme, a Bridge, a Second Theme, an optional Third Theme, and an Epilogue. The structure of the sonata form is governed by two relationships between segments: A segment may be part of another segment (signified by rhomboid arrows) and the segments are ordered in time (signified by "follows" arrows). These relations are similar to those found in a play (an act is made of scenes and the acts or scenes are ordered in time) and are captured by trees with ordered siblings like the ones that can be defined with SGML or XML document type definitions. But the relationship most interesting when analysing music is a transformation relationship: In the sonata form and in most other musical forms some segments are derived from other segments by applying a transformation. For example, in the first movement of Haydn's Symphony No. 82, the second theme is the first theme transposed to a different key. Common transformations are

- transposition of a theme from Tonic to Dominant
- expansion of a Motif
- fragmentation of a Theme
- inversion and reflection
- doubling or halving of the tempo
- changes in intensity

Many of these transformations do not operate on time intervals of the complete score but act on segments of individual voices of a piece thus complicating the correspondence between representation and realisation. In addition, as soon as more than one monophonic instrument plays, a single segment in the realisation corresponds to many segments in the representation. In essence the transformations establish relations between segments in the representation and/or segments in the realisation like inversion(A.B), i.e. segment B is an inversion of segment A. These additional relations between segments transform the structure tree of a piece of music into a structure web. Such a web can be modelled in two different ways: Either in an object-oriented flavour by making segment the central data type encapsulating all its links (cf. Fig 3) or by taking a document structure approach where the parent-child relationship is singled out to build a structure tree and the transformational links are added to the tree nodes as attributes. In both approaches the links between the segments are the dominant features to describe the representation.

The approach to architecture or document structure illustrated in Fig. 3 is prescriptive: As in 19th century music theory, it is assumed that there is an ideal architecture for a sonata form, and that a piece of music not conforming to these rules is in error. In our century this concept has come under considerable criticism (e.g. [15], [9]) and a descriptive approach has been advocated. In a descriptive approach a set of features induces a document structure that may or may not correspond to last century's idealistic conception. As a consequence, each musical work may have a plenitude of valid structures describing different aspects of its analysis or interpretation. Without the means to link these different structures to each other - and to the realisation – such a multi-faceted description is rather useless. Like with different transformational relations within one description there are two basic alternatives to structure this set of document structures: One may single out one document structure as a master and try to relate all other document structures to this master structure. Such an architecture is rather well behaved from a computational point of view; each new structure needs only to be synchronised with the master structure. The disadvantage is that comparisons between child-structures are only possible by comparing each child with the master and further processing of the results of these comparisons. As an alternative one may create a web (or net) of different structures, relating each individual structure to all others. Here all structures are peers and direct comparison between them is straightforward but when an additional structure is added it has to be synchronised



Figure 3: The prescriptive structure of a sonata form. Building blocks are links and segments. Segments have features like rhythm and harmony, serve as linkends, and can be typed with semantic labels like Motif or Particle.

with all the other structures. For more complex documents like a classical symphony, the size of the data needed to synchronise the different structures may easily surpass the size of the structure descriptions themselves.

4. TRADITIONAL "METADATA"

Besides relations within and between structural elements of a representation or a realisation, traditional metadata like author information can be interpreted as a relation (e.g. isAuthorOf(person, opus)) and therefore as a link. For minimal metadata sets like e.g. Dublin Core [3] where elements like "Creator" or "Contributor" are usually quite small and embedded in the essence as tags such an approach might be considered unnecessary complicated. But it offers at least one important advantage: It separates the description from the essence and an archive can provide *one* curated set of e.g. author data that can be linked to all pieces produced by this author. Trivial as this may seem it nicely solves the problem of spelling variants or misspelled entries generating phantom persons, a major problem in all existing catalogues. For more complex metadata sets like MIDAS

[6] a separation of description and essence becomes mandatory. Whereas in Dublin Core the creator information is just a character string, the MIDAS description scheme for an artist (Fig. 4) is a tree with a considerable size. Even a "simple" descriptor like the name of an artist has eleven elements most of which can be instantiated multiple times. A system that can make use of this complexity can find "Some like it hot" as an answer to a query for all the movies with Norma Jean Mortenson acting. But the size of a fully instantiated MIDAS description schema is definitely beyond what one would want to embed into a low bandwidth stream for annotation. Links into the tree together with a tree transformation language like XSL [1] can be used to filter specific elements from the artist description and provide the means to adapt the scope of the knowledge base to the desired usage without changing the knowledge base.

5. LINKS

Up to now the term "link" has been used as a concept but the actual linking mechanism that connects the different representations and realisations has not been specified. A



Figure 4: An person according to MIDAS

versatile linking mechanism has to fulfil a diverse set of requirements, especially it should support:

• arbitrary formats

Representations may be unstructured (plain text) or structured using a variety of markup languages ranging from standardised (e.g. SGML) over de facto standards (e.g. IAT_EX) to proprietary formats (e.g. Word). Realisations are usually stored as bit streams, nonlinear compression formats with varying bitrates are quite common (e.g. MP3), and new encoding schemes are emerging at a steady rate (e.g. Ogg Vorbis [12]),

• read only media

More and more content is distributed on read only media like CD or DVD. Even if the data can be transferred to a writeable medium, many bitstreams do not support the insertion of additional data at arbitrary locations and a link must be able to locate its linkend without modifying the data. In addition data stores for metadata like an author database usually do not grant write access to their users and therefore have to be accessed without changing the content of the data store.

• many-to-many links

Often different versions of the same realisation are available: In the case of images or music many archives store the same realisation in multiple versions with different quality (compressed for free preview and linear for fee). A data collection that is focused on performing history is likely to link equivalent segments in different realisations. And a descriptive approach to musical form has to deal with more than one representation for the same realisation.

• bidirectional link traversal

This is a direct consequence of many-to-many links. If all the linkends are peers there must be a way to traverse from each linkend to all the other ones.

points and intervals

Since the links are used to map structural elements like the scene of a play on a bitstream there must be means to address intervals in the realisation.

An immediate consequence of the requirement to address read only media is that the link information has to be self contained such that it can be stored in a separate document. The requirement to support arbitrary formats has as a consequence that the linking mechanism can not rely on its targets to supply some unique identification to specify the anchor point of the link. This problem can be solved by isolating the link from the linkend with a locator layer that deals with media-specific addressing issues and presents a unique identification for the linkends to the link proper. This indirection approach has been pioneered by the Hy-Time ([2], [7]) independent link (Fig. 5) and the mechanics of linking will be discussed using the HyTime ilink as an example. Some aspects of the ilink functionality have influenced the design of XLink, XPath, and XPointer and can be expressed in XML but its ability to refer to segments of non-SGML data is still unique. A HyTime independent link



Figure 5: Elements of an independent hyperlink

does not need to reside at any of its endpoints, the ilink tag has an attribute linkends which may specify any number of identifiers each of which points to a locator.

<myLink linkends="repA1 repA2 repA3 relA1 relA2">

One of the most useful locators is a treelocator, HyTime's ancestor of XML's extended pointer concept, which allows the navigation of untagged trees:

<treeloc id="repA1" locsrc=sgmllink>1 2 3 2 </treeloc>

Each node in the tree is identified by a list of integer values. The list of integers describes how to get from the root of tree to the specific node. The root node is assigned the number '1' and each successive number describes the position of the node among the children of the parent node by counting the children from left to right starting with '1' for the first child. Thus when applied to the excerpt from Hamlet used above, the tree locator 1 2 3 2 selects the first spoken words "Who's there":

1 specifies the root node ACT

1 2 specifies the SCENE, node 1 1 is the TITLE of the ACT

 $1\ 2\ 3$ specifies the SPEECH, node $1\ 2\ 1$ is the TITLE, node $1\ 2\ 2$ STAGEDIR

 $1\ 2\ 3\ {\bf 2}\,$ specifies the LINE, node $1\ 2\ 3\ 1$ is the SPEAKER

Since they do not rely on the presence of tags, tree locators provide a powerful addressing mechanism that covers a wide range of special cases like e.g. lattice structures or matrices. Whereas HyTime does not predefine the locators needed to refer to segments in the realisation it provides mechanisms to define coordinate systems and measurement units which can be used to define "custom-made" locators (for details see [7, 2]) for points and intervals in arbitrary media.



Figure 6: Linking Hamlet

Inverse traversal from anchor to link is possible by searching for the locators that cover the current position and selecting all the links referring to these locators.

6. GENERATING THE LINKS

In order to link a performance of Hamlet recorded on CD with the tagged text of the play a separate link document has to be generated that connects recorded utterances with the text and imprints the structure information on the audio stream (Fig. 6).

To simplify the description of the link generation (Fig. 7) it is assumed that a linear recording of the performance (e.g. a WAVE file) and an SGML or XML tagged text of the play are available. The principle processing steps are illustrated in Fig. 7: Structure and plain text are extracted from the tagged representation and separated. The plain text is decorated with time tags that specify start time and end time of each word in the plain text. This timestamped representation is merged with the extracted structure and formatted conforming to HyTime syntax.

Generating the plain text from the tagged Hamlet representation requires some pre-processing: All tags that markup unspoken content (ACT, SCENE, etc.) and all implied markup like punctuation are filtered from the representation and from the remaining tags only the untagged content



Figure 7: Generating the links

is used:

Who's there Nay answer stand and unfold me He vourself Long live the king Bernardo You come most carefully upon your hour

The structure is represented by a sequence of tree locators for the spoken words indexed with the words:

1	2	3	2	1	Who's
1	2	3	2	2	there
1	2	4	2	1	Nay
1	2	4	2	2	answer
1	2	4	2	3	me
1	2	4	2	4	stand
1	2	4	2	5	and
1	2	4	2	6	unfold
1	2	4	2	7	yourself
1	2	5	2	1	Long
1	2	5	2	2	live
1	2	5	2	3	the
1	2	5	2	4	king
1	2	6	2	1	Bernardo
1	2	7	2	1	He
1	2	8	2	1	You
1	2	8	2	2	come
1	2	8	2	3	most
1	2	8	2	4	carefully
1	2	8	2	5	upon
1	2	8	2	6	your
1	2	8	2	7	hour

The audio recording is fed through a speech recognition engine that produces a transcript and tags each recognised word with its start and end time. As can be seen from Fig. 8 [19] there is an n:m relation between the plain text (called reference text in the figure) and the speech recognition transcript. Plain text and transcript are aligned with a dynamic programming algorithm [4] and after alignment the word times from the recognised text are transferred to the corresponding words in the plain text. Most of the speech recognition errors do not influence the timing values: An isolated word recognition error (alignment quality 1 in Fig. 8) still gives the correct timing. An error were instead of a single word in the representation a group of words is recognised (alignment quality 2) does not pose a problem, the correct word is marked with the start time of the first word and the end time of last word in the recognised group. Only errors where a group of words in the representation is transcribed as a group of incorrect words (alignment quality 3) yield unrecoverable timestamp errors. In this case the missing time values can either be generated by interpolation or the timing for the afflicted word boundaries can be marked as unknown. After aligning and resolving the speech recognition errors we can tag the words from the plain text with start and end times:

Who's	210	211
there	211	215
Nay	220	221



Figure 8: Aligning plain representation and the transcript

221	226
227	228
232	234
235	236
237	242
242	250
260	262
262	263
264	266
266	267
275	276
280	281
290	292
292	294
294	296
297	301
301	302
303	304
304	305
	221 227 232 235 237 242 260 262 264 266 275 280 290 292 294 297 301 303 304

This time tagged plain representation is merged with the word tagged structure

T	2	3	2	1	Who's	210	211
1	2	3	2	2	there	211	215
1	2	4	2	1	Nay	220	221
1	2	4	2	2	answer	221	226
1	2	4	2	3	me	227	228
1	2	4	2	4	stand	232	234
1	2	4	2	5	and	235	236
1	2	4	2	6	unfold	237	242
1	2	4	2	7	yourself	242	250
1	2	5	2	1	Long	260	262
1	2	5	2	2	live	262	263
1	2	5	2	3	the	264	266
1	2	5	2	4	king	266	267
1	2	6	2	1	Bernardo	275	276
1	2	7	2	1	He	280	281
1	2	8	2	1	You	290	292
1	2	8	2	2	come	292	294
1	2	8	2	3	most	294	296
1	2	8	2	4	carefully	297	301
1	2	8	2	5	upon	301	302
1	2	8	2	6	your	303	304
1	2	8	2	7	hour	304	305

and after eliminating the words the result is a raw link file connecting the representation with the realisation

1	2	3	2	1	210	211
1	2	3	2	2	211	215
1	2	4	2	1	220	221
1	2	4	2	2	221	226
1	2	4	2	3	227	228
1	2	4	2	4	232	234
1	2	4	2	5	235	236
1	2	4	2	6	237	242
1	2	4	2	7	242	250
1	2	5	2	1	260	262
1	2	5	2	2	262	263
1	2	5	2	3	264	266

1	2	5	2	4	266	267
1	2	6	2	1	275	276
1	2	7	2	1	280	281
1	2	8	2	1	290	292
1	2	8	2	2	292	294
1	2	8	2	3	294	296
1	2	8	2	4	297	301
1	2	8	2	5	301	302
1	2	8	2	6	303	304
1	2	8	2	7	304	305

that needs only reformatting to become a valid HyTime hub document.

7. CONCLUSION

Independent hyperlinks are a versatile and powerful mechanism to link between and among representations and realisations and to establish relations between data. Since the linkends can be arbitrary objects, nothing precludes their interpretation as object and interpretant and thus the link becomes a sign (or the sign a link): "A Sign, or Representamen, is a First which stands in such a genuine triadic relation to a Second, called its Object, as to be capable of determining a Third, called its Interpretant, to assume the same triadic realation to its Object in which it stands itself to the same Object."[14], quoted from [5]. Adopting this interpretation allows to rephrase the title of the paper to something more fitting for a semiotics conference: "The Sign is the Data".

8. REFERENCES

- J. Clark. XSL transformations (XSLT) 1.0. Recommendation, World Wide Web Consortium, Nov 1999.
- S. J. DeRose and D. D. Durand. Making Hypermedia Work - A User's Guide to HyTime. Kluwer, 1994.
- [3] Dublin Core Metadata Initiative. Dublin core metadata element set (DCMES) Version 1.1. Recommendation, Dublin Core Metadata Initiative, 1999.
- [4] D. Gusfield. Algorithms on Strings, Trees and Sequences. CUP, 1997.
- [5] C. R. Hausmann. Charles S. Peirce's Evolutionary Philosophy. CUP, 1993.
- [6] L. Heusinger. Marburger Informations-, Dokumentations- und Administrations-System (MIDAS). K. G. Saur, second edition, 1992.
- [7] ISO/IEC JTC1/SC18/WG8. Hypermedia/time-based structuring language. Technical report, 1997.
- [8] L. Jin and Z. Wen. Adorning VRML worlds with environmental aspcts. *IEEE Computer Graphics and Applications*, 21(1):6–9, 2001.
- [9] C. Kühn. Form. MGG Sachteil, 3:607-643, 1995.
- [10] F. Long, D. Feng, H. Peng, and W.-C. Siu. Extracting semantic video objects. *IEEE Computer Graphics and Applications*, 21(1):48–55, 2001.

- [11] S. Mauser, editor. Handbuch der musikalischen Gattungen. Laaber, 1993–.
- [12] http://xiph.org/ogg/vorbis/index.html. Technical report.
- [13] G. Papaioannou, E.-A. Karabassi, and T. Theoharis. Virtual archaeologist: Assembling the past. *IEEE Computer Graphics and Applications*, 21(2):53–59, 2001.
- [14] C. S. Peirce. 1902 manuscript.
- [15] C. Rosen. Sonata Forms. Norton, 1988.
- [16] A. Sheth and W. Klas, editors. Multimedia Data Management. McGraw Hill, 1998.
- [17] I. Shlyakhter, M. Rozenoer, J. Dorsey, and S. Teller. Reconstructing 3D tree models from instrumented photographs. *IEEE Computer Graphics and Applications*, 21(3):53–61, 2001.
- [18] C. Sperber-McQueen and L. Burnard, editors. *Guidelines for Electronic Text Encoding and Interchange*. 1999.
- [19] J. Stöffler. Optimizing speech recognition resources for text-audio matching, 1998.

the data][h!][bleeding/m[ez]ang.elle.d texts: Constructing Polysemic & Neology Fic/Factions Online

Mez [Mary-Anne] Breeze netwurker@hotkey.net.au

1)Ev.o[h!]lution::Pre Alphanumeric//Mezangelleing Daze::

if:

pre alphanumeric//pre network n-cluded use ov com.put[ty/fillah]ers offline

then:

n-turr-rest in network system [ic]z stemmed fromme a more organic base, collaborationz were via real-time flesh meat N n-stallation based

[d-tail n-hance.ment//philo.soph[omore]icall link wish n-sert:

movin amongst the pack[ette] ov life, we n-gage/ah-lert senzez that latah pull threads 2gether in2 a netwurked w.hole. xsamp.elle::when present-ting a papah on comp[lex]uter art in 93, the seed waz so[?]wn N x-periment.all edgez merged in2 the format of the aware, the concreted creatiff shift. all i hadde 2 do waz wait]

then:

the phirst html do[a]cumen.t <u>cutting spacez</u> gestation waz in 1995, a hybridity year encased via a terminal discovered on the campus of wollongong university. many h[ertz]ourz spent chatting on Cybersite and <u>Kajplats</u>, channels <u>dis[cobolos]</u>coverin the er[k]rrorz N <u>jo[uissance]</u>y of this new textual realm. jumping fromme one terminal 2 the next//running three chat-roomz at once via three diffurrent terminal[behaviour]z sew as 2 opt.tim[id no lonah]ize the time d-layz, chairz blurring b-tween as the monitorz flashed fiction wurdz that [k]needed...bleed.ed e.vent[ingz]ualli into the cutting spaces docco...

[pause]

brief 4ray int MOOs; e.mail dependancee startz; transfer.all ov play became crucial here, while avatar patterns b-came e-veedent/m-merg-NT. di-wreck MOO based e.moti[vate]ng ncorporated in2 general fiction cre[che]ation. every.da[ze]y chrono.logical time constricturez erased; temporal awarenezz irretrivably alt[.ctrl.delete]ered; geophysical statez became a hazy con[ned]struckt ov a "rl" [real life] parameter, with text no longah static and punct.ewe.ation actualizing in2 langue.

[d-tail n-hance.ment//physical link wish n-sert:

the computer became the real. terminal waz no longer a wurd 2 b feared, n-dicating abortive allusionz, but N actual object dsign[n symbol]ed 2 n-hance, open, mutate. _my_ time became _mi_ own, mi text, mi fictional boundaries gellin with.in otherz azz yet known in a physical senze. blood N bone where no longah paramount outside of a textual wreck.cog[ging][ig]nition. B-ing parte ov the network opened up the fiction gatez N eye crossed ova in2 Bled Text:

[see the first won; _cutting spaces_ captured thizz s-sence of thredding and bleeding spaces on & offline, making the line bend N x-tending in2 both functional states. Oth[ub]er projects pre/sent-mezangelle n-clood _a disgruntled book of wizzdumb_//_Blood Puppetz_// Project M-E-Z// etc ad nauseum....

cyberspace b-came siphonspace, cypherspace, psychastheniaspace. Mz Post Modemizm's birthe pre-M-ted her knowledge ov avatarian statez. 96 bought aff.[terra]firma.tion; mailing lizt act.ion became de rigueur; until 98 the projectz flowed N the mails spreade bac N 4^{th} , 4^{th} N bac...ensemble///fleshfactor//7-11//American

Express//recode//rhizome//nettime - u name it, avartarian statements where there. avatarz breeding in this trubbled space where mani, with psiborgs b-cumming sciborgs b-come.in[here] <u>sighborgz</u>....

[if mez =

mezchine//emauler//mz

post mortemism/modemism//flesque//mezflesque.exe//etc

then i.d.entitee = flow = <u>internal damage report</u> & <u>fleshistics</u>]

[deep x-posure 2 telesthesia & teledildonics allowed fictiothermal currents 2 circ[e].u.late; collab.awe.ration grew frome others in the stream respon.d[ance]ing, playin, without reference 2 any narratiff structure ore awareness ov procedure]

wot a time.

2)[r]Ev.o[h!]lution::Alpha+betanumeric M[aging]ezangelleing Daze::

if:

Brett Stalbaum asked:

>There is a REFRESHING laconic quality in your net.art, especially given that

>you participate primarily in the rubric of hypertext poetry. This extends not

>only to the focused verse, but also to your consciously navigable, clear

>hypertexts. What are your thoughts on the prolix vs. the laconic in net.art

>hypertext?

then:

[I anne.swered:]

<Obviously the medium of hypertext lends itself readily to the minimal....minimal in terms of a primary reliance the most basic elements - text, screen ['doc swapping'] and image. My laconic use of these most base.hic[!] elements [or elle-E.ments as i would say if responding via regular mez/nschine communication channels] is governed largely the need to condense/dilute/refresh wordage and imagery meanings/established codes/cues of associative meaning/s.

<Hypertexting has pow!er, and a reader can n-gage this through x-tremely personalized connections/no[h]des if given the opportunity...if a reader is presented with a tract of informatically direct/explanatory dis[ease]course [whether it be fiction, faction, theory, etc] a passive absorption interaction channel is more likely to result. If an intr.ACT[or]a [or reader in preMezANGLEtext] has to cog[wheel]nitively spark with the work presented, and does not have access to s-tablished lingual and pictorial threads planned out and ready to be spoon/key fed, imagine the result! Confusion, granted, but if an int.ACTa can draw some webthread, or nuance, from the wurr.k, then.....[insert random jouissance here:)]

<In short, the KISS princ[ess]able has some me[z]rit:

<Keep It Simple Stupid.

<In my case, it should read:

<Keep it [conceptual and] Sizzlingly Short

[d-tail n-hance.ment//e.lab.ore.ation link wish n-sert:

now came the merg.her, merge.awe, merg.whore. technique b-came theory; fiction fact N the dis.<u>tinct[ure]</u>ion e-rrelevant. emailing turned fromme fictional struc.t[o]ur.ez

2 cracking otherz wurdz, re-alignin them, reversin, refemmeing in2 a medical/sci-fi mash; mixing theor[.M.z]ee in2 postfiction, faction]

[plunge]

2 mezangelle izz 2 now take back the fiction koda and twizt//reroute a narratiff thru warped be.lie.f system sy.[e]mbolizm [religion:<u>the DataH Inpho[mill]ennium_]</u> N [cloning tech.no.logee:<u>The Clone Alpha Project_]</u>. Techniques shifted fromme the struck.tor[tor]aul 2 dynamizm [flashed up or <u>director based</u>] but the textual L-ee.ment remained the same.

[datahbleede.r[& inge] + o[h!]pen source koda spawnz. collaboratorz swoop. text re-inne.ventz. translationz spark n sticke 4 a l'tle while.]

3)[e]Vol[ve]ition::Omeganumeric Mezangelleing::

2 mezangelle::

2 4m a text fromme the ground[ing] uppe

2 n-hance the simple text of an email thru the splicing of wurds

2 phone.tic[k-tock]aulli m-bellish a tract ov text in2 a neo.logistic maze

2 network 2 the hilt N create de[e]pen.den[ting]cies on email lizts for the wurkz dis.purse.all

2 graphi.caulli N text.u.alli e-voke a conscious sens.u.all & lingual mix

2 make net.wurkz space themz.elves in2 a spindle of liztz thru collaboratori n-tent

2 uze computer kode kon.[e]vent.ionz spliced with irc emoticons and ab[scess]breviations

2 spout punctu[rez]ationz reappropraited in2 sentence schematics

- 2 polysemicalli m-ploy a fractured wurdset
- 2 m-brace 4m conventionz
- 2 flaunt pol[emic]itical l-usions
- 2 ig.gnaw word endinz
- 2 let lettahs b used as subsetz
- 2 x tend N promote n-clusive meaningz.

... Seine hohle Form... - Artistic Collaboration in an Interactive Dance and Music Performance Environment

Joseph Butch Rovan CEMI – Center for Experimental Music and Intermedia University of North Texas, PO Box 13887 Denton, TX 76203-6887 +1 940 565-3762

rovan@music.unt.edu

ABSTRACT

In this paper we outline some of the technical and aesthetic concerns involved when musicians and choreographers collaborate within the context of an interactive dance system. Issues of mapping are examined, focusing on the mapping of dance gesture to real-time synthesis parameters. Perceptual correlation of these mapping strategies is stressed, albeit through varying levels of abstraction. Finally, portions of the authors' interactive computer music/dance work "Seine hohle Form" are profiled as examples of the aesthetic and technical challenges faced when working with such systems.

Keywords

Dance, interactive dance system, gesture, interactive computer music

1. Introduction

The use of choreographic gesture as a control component in music composition/performance for dance has been a concern of choreographers and musicians for almost half a century. As electronic instrument builders of the 20th century struggled to devise unique interfaces for their unique instruments, choreographers such as Merce Cunningham offered the surprising option of extending the concept of gestural control to the world of dance. The Cage/Cunningham experiments of the 1960s using Theremin technology to sense body motion are only one example of this experiment that still continues today.

First published at CO SIGN-2001, 10-12 September 2001, CWI, The Netherlands. Robert Wechsler / Frieder Weiss PALIDROME Inter-Media Performance Group Johannisstr. 42, 90419 Nürnberg +49 911 39 74 72 {robert,frieder}@palindrome.de

When musical control was relinquished to dance gesture, the union of open-air (non-contact) gesture to sound raised many intriguing questions. Even though the technology has progressed to the point where current dance systems rely on sophisticated video tracking instead of the antennae of a Theremin, the cause and effect relationship between sound and gesture has remained an elusive problem. Up to this day, most interactive dance/music systems have relied on fairly simple relationships between gesture and sound, such as the basic presence or absence of sound, volume control, possibly pitch control.

The lack of progress in this direction has been complicated by the tenuous threads of communication between the computer music and dance fields. Indeed, although much work has been done recently in the world of computer music by composers/performers developing and composing for gestural controllers, the world of dance has remained largely isolated from these developments.

Today's tools, however, provide the possibility of rich relationships between dance and music in interactive systems. Real-time software for music synthesis and digital signal processing (i.e., MAX/MSP, developed by Miller Puckette and David Zicarelli, and jMAX, developed at IRCAM) is readily available and runs on standard desktop and laptop PCs (Macintosh and PC LINUX). Likewise, comparable developments in video image tracking/processing as a source of gestural information (e.g. Palindrome's EyeCon system) have given composers and choreographers powerful tools with which to harness the expressive gestures of dance. But the lack of communication between the two fields, and the often-limited concept of what interaction actually means in such work has limited, in the authors' opinions, the expressive possibilities of the collaborative work.

Working alternately in Nürnberg, Germany, and Denton, Texas, Palindrome Inter-media Performance Group and the Center for Experimental Music and Intermedia (CEMI) have explored these issues in their ongoing work together. A body of interactive dance/computer music works is emerging, as well as a dancespecific vocabulary of gesture mappings between movementrecognition and real-time digital sound synthesis.

2. Mapping

In an interactive system, sensors do the job of "translating" one form of energy into another. Specifically, the physical gestures of dance are translated via sensors of various sorts into a digital signal representation inside of a computer. One the gesture is available as an abstract amount of computer data, however, the important question arises: what do we do with it?

"Mapping" is the process of connecting one data port to another, somewhat like the early telephone operator patch bays. In our case mapping has a very specific connotation—it means the applying of gestural data, obtained via a sensor system, to the control of some sound synthesis parameter. The dramatic effectiveness of a dance, however, invariably depends on myriad factors—movement dynamics of body parts and torso, movement in space, location on stage, direction of focus, use of weight, muscle tension, and so on. And although sensors may be available to detect all of these parameters, there still remains the question of which ones to use in a given setting, and then to which of the equally many musical parameters to assign it.

Herein lies the basic quandary. Making these mapping choices, it turns out, is anything but trivial. Indeed, designing an interactive system is somewhat of a paradox: the system should have components (dance input, musical output) that are obviously autonomous, but which, at the same time, must show a degree of cause-and-effect that creates a "perceptual" interaction. Unless the mapping choices are made with considerable care, the musical composition and choreography can easily end up being slaves to the system. In some cases, interaction might not occur at all. Not in a technical sense—the movement will indeed control the music—but in the sense that no one (except perhaps the performers) will notice that anything is going on!

Some may argue that it doesn't matter whether or not an audience is aware that interaction is taking place. Even if the artist is completely alone in experiencing the interactivity, for some it may be enough that the system of interaction "privately" affects the performer's expression within the piece. The audience is only vicariously part of the interactive experience.

Palindrome Inter-Media Performance Group has not chosen approach. Instead, we have searched for a more effective—or at least more convincing—result. Two strategies seem reasonable. One is, as mentioned above, for the choreographer and composer to create their work specifically for a given technological system. Not, of course, that every dance gesture needs to trigger every musical event—there is actually considerable playroom in this regard. What Palindrome has learned is that even when only *part* of an interactive piece is clear and convincing, audiences will attune to and accept more complex relationships.

The second strategy, which does not exclude the first, entails more deliberate and targeted mapping strategies. This is in turn means identifying those qualities or parameters which are most fundamental to a piece—analyzing what information lies in the essential nature of the work. This is a more complicated, but rewarding approach, since it means that the technical system is born out of a need to serve the artistic vision, instead of the other way around. These mapping strategies should focus and harness the essential parameters of the movement, while acknowledging the perceptual aspects of human movement. A first step toward this approach might be to begin with the relationship of that which our eye sees, to that which we perceive when watching human bodies in motion.

3. Gestural Coherence

Just as is true of the sound world, we do not perceive the human body in motion in a very objective or scientific way. What we perceive in dance is highly filtered and often illusory--the choreographer and dancer work hard for this effect. Indeed, the quality of flow at one moment may dominate our perception of a phrase so much so that individual shapes of the body go unnoticed. At another moment, geometrical shapes may override our perception of movement in space, and so on. And of course the sound--particularly musical sound--has a powerful affect on how one perceives dance.

Our projects in Palindrome have explored these issues of perception and movement. In particular, we have concentrated on the notion of "gestural coherence"; that is, the perceptual coherence between sound and the movement that generates it. Within the context of this search for gestural coherence, we have made the following postulations:

- An emergent integrity arises when the relationship between the dance and music systems is "believable".
- Believability depends upon gestural coherence.
- Gestural coherence is achieved through a system of mapping that mediates the two parallel structural systems (musical and choreographic).
- Musical structure emerges from dance gesture through a schema that provides for a mixture of the following gesture-to-synthesis parameter mapping strategies:
 - one-to-one, or "direct" mapping
 - one-to-many, or "divergent" mapping
 - many-to-one, or "convergent" mapping

4. Application: "Seine hohle Form"

The words "seine hohle Form" are a fragment from the poem "Gesichter" by Rainer Maria Rilke, roughly translating to "its hollow form." As a starting point for this interactive work, premiered at CEMI in November 2000, the title words serve as an emblem for the interesting challenge of creating a musical work that only exists when a dancer moves, and a dance in which movement must be approached as both functional, musiccreating gesture as well as expressive or decorative elements. The collaboration between music and dance on this piece was complete; that is, the movement and sound were not designed separately, but interactively. Indeed it could hardly have been otherwise. The choreography is affected by the live generation of sound through the use of sensors and real-time synthesis, and the resulting music is in turn shaped by these movements. There are no musical cues for the dancers, since without their movements the music is either nonexistent, or at other times, missing key elements. This method of working forced not only an inherent degree of improvisation upon the group, but also prompted a sharing of artistic roles in the working process: dancer became musician, composer became choreographer...

"Seine hohle Form" is of course not the first interactive computer-controlled dance. As mentioned earlier, interactive dance has a long history, and recent important contributions include the work of David Rokeby, Troika Ranch, Antonio Camurri, among others. Our work may be unique, however, in the extent to which multi-dimensional mapping strategies have been used within a framework of gestural coherence.

4.1 Technique

In Palindrome works the dancers' gestures are tracked using the EyeCon video-tracking system, designed by Frieder Weiß of Palindrome. EyeCon is based on frame-grabbing technology, or the capturing of video images in the computer's memory. By frame-grabbing and processing a dancer's movements, it is essentially possible to convert their gestures into computer data that can then be mapped into control of music or other media. For "Seine hohle Form" we use three small video cameras set up

above and diagonally in front of the stage.

The analysis features of the EyeCon video-tracking system include the following six movement parameters:

1. Changes in the presence or absence of a body part at a give position in space.

2. Movement dynamics, or amount of movement occurring within a defined field.

3. Position of the center of the body (or topmost, bottommost, left or rightmost part of the body) in horizontal or vertical space.

4. Relative positions (closeness to each other, etc.) of multiple dancers (using costume color-recognition).

5. Degree of right-left symmetry in the body--how similar in shape the two sides of body are.

6. Degree of expansion or contraction in the body.

While in theory these features may be combined in any way, in practice a maximum of two at one time seems to make sense; more, we have found, is simply not perceivable.

The real-time sound synthesis environment was designed in MAX/MSP by Butch Rovan. A PC running EyeCon is linked to a Macintosh PowerBook running MAX/MSP, sending the gestural data gathered by EyeCon to the real-time sound synthesis parameters. All mapping is accomplished within the MAX/MSP environment, and changes throughout the work. (fig. 1)



Figure 1: The real time sound System environment of ... seine hohle form...



Figure 2: Custom-build DSP modules of ... seine hohle form...

The MAX/MSP program for "Seine hohle Form", designed by Butch Rovan, is a musical synthesis environment that provides many control parameters, addressing a number of custom-built DSP modules that include granular sampling/synthesis, additive synthesis, spectral filtering, etc. (fig. 2)

Control of the musical score to "Seine hohle Form" is accomplished through a cue list that enables/disables different mapping and DSP modules. Both EyeCon and MAX/MSP software components are organized as a series of "scenes", which represent different configurations of video tracking, mapping, and DSP.

4.2 Examples from "Seine hohle Form"

The following description of excerpts from "Seine hohle Form" is certainly not complete; even within these few described scenes there is a good deal more going on. Nevertheless, it may offer an introduction to our working methods. (NOTE: a QuickTime movie excerpt of SHF is available at www.palindrome.de).

In scene number 5, the first dancer (female) controls nine relatively clear and isolated additive synthesis tones with the extension of her limbs into the space around her (an example of one-to-one mapping). An algorithm in MAX/MSP modifies the pitch and timbre slightly with each extension. Meanwhile, the second dancer (male), back to audience, uses small, whole-body movements to cut off quieter, "whiter" sounds which build continuously as long as he is not moving.

In scene number 1, the male dancer manipulates a stream of loud, aggressive sound fragments derived through granular sampling. He activates the sounds through equally aggressive side-to-side torso movements. The speed and velocity of his movements shape the parameters of the granular sampling engine continuously, with many interactions between incoming gesture parameters (an example of convergent mapping).

In scene number 3, the male dancer finally rises from his low stance and approaches the audience. Here, his height (highest body-part from floor) controls the parameters of a real-time spectral filter, producing a thinner and more continuous musical texture. The effect is much subtler and less direct than what has come before, and lends a sense of disorientation to his part, softening his role following the opening solo, and thus opening the way for the female dancer to begin her own solo.

5. Conclusions and Future Work

The basic technical system described in this paper has been operational for almost a year (and had been tested in performances in Munich, Dresden and, most recently at the 2001 conference of the Society for Electro-Acoustic Music in the United States in Baton Rouge, Louisiana). It has, however, become increasingly clear to us that our current process for gestural mapping could be improved by creating a clearer hierarchy among the parameters that govern relationship between the video-tracking system (EyeCon) and the sound synthesis software (MAX/MSP). In particular, we are working to segregate more clearly the tasks that are assigned to each component of the system.

Of course, making use of the inexhaustible mappings between movement and sound requires an understanding of the different and potentially conflicting-goals that drive composers and choreographers. In the past, traditional models of collaboration between composers and choreographers have subjugated either dance or music, or sidestepped the question altogether by removing all correlation between movement and sound. In a collaborative work such as "Seine hohle Form," a new opportunity exists, one that avoids this conflict entirely by making the work of choreographer and composer interdependent rather than dependent; fused instead of segregated.

6. ACKNOWLEDGMENTS

The authors would like to thank CEMI (Center for Experimental Music and Intermedia) and 01plus Institute for their assistance and support. Thanks also to Helen Zwiauer and Jon Nelson.

Streams of Motion (Super Spectacular) A Virtual Reality Art Work

Tim Portlock

MFA candidate Electronic Visualization Laboratory University of Illinois, Chicago 842 W. Taylor Street Chicago, IL 60607 U.S.A. portlock@evl.uic.edu

ABSTRACT

The Streams of Motion (Super Spectacular) project was originally created as part of a larger project consisting of several environments created by individual artists. The goal of this work is to begin a long-term exploration of certain issues particular to the process of creating art within the medium of virtual reality. Specifically, what types of creative processes could be developed that would allow the most direct way of going from concept to realization. This exploration in turn effected the how I approached subject matter, content and narrative structure of the project.

The work was designed to be experienced in the CAVE, a four screen, rear projected, immersive virtual reality system, but may be adapted to a smaller format virtual reality display device.

KEYWORDS

Virtual Reality, Art

1. INTRODUCTION

While the hardware necessary to make virtual reality is rapidly becoming more accessible, the knowledge required to take advantage of these tools are not yet as readily available. This is an especially significant issue in that most forms of virtual reality typically require

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

any number of a wide range of skill sets; programming, computer modeling, image processing, audio engineering, etc. Because of the range of skills involved there are certain preexisting artistic production models that are most applicable to developing a virtual reality project.

1.1.1 The Film Production Model

By "film production model, I mean a group of people performing separate functions on a single project. An additional benefit deriving from this model is more work can be done in less time and the capacity to build larger and or more complex environments is increased.

1.1.2 The Solo Artist Model

In the "solo artist model, a single artist either works on a single complex virtual environment or on several relatively smaller and less complex virtual environments.

1.2 The Solo Artist Model in the Digital Environment

Because working from the Film Production model is less feasible outside of the institutional setting I have begun to experiment with ways of working within the Solo Artist production model that enables me to create complex virtual experiences. This was one of the issues I chose to explore through the "Streams in Motion" piece. One of the main things I wanted to accomplish in "Streams" was to begin a long-term exploration into ways of working that would allow a single artist to develop a rich and complex virtual environment.

2. WORKING PROCESS 2.1 Technical Aspect

A technique that has emerged from computer games is to suggest complexity without actually articulating the correlative details. The standard approach in the computer game industry is to use textures and models several times in one game. Reuse of elements saves processing resources, which improves the quality of the game's performance. An added benefit of reuse is that it economizes the amount of labor required to develop a project. The important thing is to reuse elements in such a way that their reoccurrence does not feel repetitious to the user.

Drawing on my background in painting, the second technical step I took to economize time was to handdraw a large portion of the textures used in "Streams." While I am primarily interested in using hand-drawn textures in this virtual/digital environment for their narrative meaning, this approach also streamlines the overall development process. On previous projects a significant portion of the overall labor was spent processing textures with applications like Adobe Photoshop. Drawing the textures enabled me to imply volume and detail with an economy of effort. Because the images are black and white I required less time to process them.

2.2 Narrative Aspects

Because of the relative newness of virtual reality as an art form there is not yet a typical approach to creating and presenting narrative. Artists that choose to create non-linear narratives have to develop ways to create tension within a story that can evolve in any number of directions and still remain engaging to the user. Ideally the development of a non-linear narrative involves some level of testing to determine how user interactions effects the evolution of the story line. This can be as informal as the responses of friends.

The narrative in my project was created modularly. Each module contains an architectural structure and sets of interactive events. These modules can be experienced as stand alone narratives or they can be placed near each other in a single virtual environment and, if routed, be experienced as a single narrative.

Two of the major advantages of working modularly are that the order in which the story branches can be reconfigured with relative ease and that the artist can develop the narrative in manageable portions.

2.2.1 Narrative Modules 2.2.1.1 The Starting Point

This is literally where the user first enters the Virtual Environment (VE). Although the starting point could have been set anywhere in the landscape, I chose to start the user out in a parking lot, an area associated with embarking and disembarking. (see figure 1.)





2.2.1.2 The Amphitheater

The user enters the amphitheater lobby. In the center and on both sides of the lobby are concession stands each selling a different set of products; toys, snacks and glue guns each produced by the same company sponsoring the event. The glue guns resemble real guns. Some of the items at the concession stands are "grabable" by the user.

By making this object available I am attempting to set up expectations in users about how to behave in the environment based on past experience with first person shooter games. It is these types of expectations and the surprise generated by thwarting them that I use to develop some of the narrative branches.





Beyond the concession stands are the entrances to the different tiers of seating for the amphitheater; ringside, first level and upper level. Each level of seating reflects the differences in ticket pricing (see figure 2). In the center of the lower most level of the amphitheater is the ring, which contains two robotic combatants each visually resembling the standard hero and villain of the 1970's Blaxploitation film genre. (See figure 3.)



Figure 3.

2.2.1.3 The Factory

The Factory consists of two main components: the section that contains the industrial machinery and the section that contains the horse coral. Users enter the factory through the front office where they are offered an automated tour of the interior of the building. The fully automated factory also contains museum type displays of historical interest. While on the tour users are given the opportunity to operate the main controls of the factory, which can alternately produce glue guns, snack foods or toy horses. (See figure 4.)

parked cars, weeds and the footprints trails left behind by each of previous users. If users get lost in the exterior landscape, they can retrace the trail that they have generated. The footprint trails are also artifacts of

how previous users have interacted with the virtual environment. They may suggest to a user a particular route through the environment, which might encourage him or her to follow suite and thus to experience the landscape in a particular order

3. CONCLUSION

For several hundred years Western image making has been concerned with the modeling of space. In virtual reality, computers create the space the artist must work within.

The medium presents a particular space, a one point linear perspective, a model that has a set of associations and meaning that I am interested in reexamining in light of it's default presence in virtual reality.

4. ACKNOWLEDGMENTS

My thanks to Jim Sosnoski for editing.



Figure 4.

2.2.1.4 The Nuclear Power Plant

Balloons and signs on the exterior indicate to the user that the power plant is holding an open house. Users enter through the front door where an employee aquatints them with the equipment used to operate the nuclear reactor. The visitors and user are given the opportunity to experiment with the controls of the power plant.

2.2.1.5 The exterior landscape.

The entire project is black and white. Sky and ground are both white making it impossible to discern a horizon line. The only visual elements assuring users that they are indeed walking on a flat plain as opposed to a white void are the various architectural structures,

Literary Theory and Computer Games

Julian Kücklich LMU München Brudermühlstr. 28 81371 München ++49 89 726 26 808 kuecklich@gmx.de

ABSTRACT

In this paper I discuss the possibilities and limitations of analyzing computer games from a literary perspective. Starting from a critique of the 'theoretical imperialism' of most ventures so far to use philological terminology in the study of computer games I attempt to assess the merits of this perspective and its contributions to a general theory of interactive fiction. While I am mostly concerned with narratological aspects of computer games, I also try to define areas of inquiry for which the terminology of literary theory is not appropriate.

Keywords

Literary Theory, Computer Games, Narratology, Aesthetics

1. INTRODUCTION

In evolving toward an integrated science of cultural phenomena and the media, literary studies have turned toward new fields of analysis. These fields now include not only literature in all of its different forms and varieties, but also films, hypertexts, and art forms that explore the possibilities of computer and video technology. However, the analysis of these phenomena remains dominated by the paradigm of the printed text – and although the term 'text' has come to signify an increasing number of things, artifacts such as computer games are still being neglected by literary studies. In assuming that lingual and scriptural signs play only a marginal role in these phenomena, this perspective disregards that the processing of signs always recurs to language in some way - even understanding a picture, or a sculpture, requires some sort of literacy to decipher the object's references to cultural codes. And, as Nicholas Montfort puts it, "even in a purely graphical interactive fiction the interactor must do some internal reading as he or she pieces together the narrative from the images displayed. This is akin to the non-verbal 'reading' done by someone looking at a picture book or a narrative series of photographs."[1]

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands.

2. TEXT AND CODE

Under these preconditions, I consider it legitimate to regard computer games as texts; yet it is impossible to predict if such a perspective will yield meaningful results. There are several points which make it seem worthwhile to approach the field of computer games from a literary perspective. For one, many computer games are based on a literary genre such as the spy novel. But even more important is the fact that there is a plot to many computer games; a narrative element that most traditional games lack. Furthermore, literary studies have embraced the metaphor of the game as a means to describe the constitution of a text from the complementary acts of writing and reading.

If there is an argument against regarding computer games from a literary perspective, it is of epistemological nature. Due to its interactivity, the object of analysis is prone to change according to the way it is approached. While this problem is certainly encountered when reading a printed text, within a computer game it becomes almost impossible to differentiate between manipulations of the objective text and its subjective actualization, i.e. between text and reading. This dilemma can only be resolved by shifting our attention to a more profound plane, that is by regarding the program code as the actual text of a computer game. We can then define the text as a set of rules that governs the fictional world of the game, whereas the game itself is merely an individual reading of this text. This model also explains why playing a computer game will never actually be the same experience for two different individuals. The confusion of these two planes of the computer game mainly results in regarding the individual game as a text whose signs must be deciphered. Peter Bøgh Andersen's semiotic approach to computer games [2], for example, is doomed to fail because he interprets the signs on the interface at face value. Semiotics has not yet supplied us with a model that would enable us to classify the different ways in which the rules of the code become manifest on the screen; although it might be rewarding to modify Charles S. Peirce's model of the abductive calculus for this purpose.

Dalum and Sørensen's narratological analysis of the sciencefiction game *Buried in Time*, is, on the other hand, an attempt to take individual readings into account by regarding them as different versions of the same story. While this is a promising approach, it still lacks the perspective needed to look beyond what is happening on the screen. Apparently, the authors themselves regard their terminology as inadequate, since they concede that "even though the 'newness' of the interactive media may, for the time being, warrant [a traditional approach] we still feel that Friedman is right in suggesting that a full description of the new media does in fact entail the development of new theoretical approaches to understanding the computer as a medium." [3] This lack of perspective is typical of the literary scholar's approach to computer games. Transfixed by the signs on the screen, she tends to forget the code behind the interface. It is important to keep in mind, however, that this code can be manipulated, while the interpretation of a traditional printed text only changes it at a superficial level. And this manipulation goes beyond the possibilities for interaction offered by the interface. Computer 'gamers' are obviously more skilled at deducing the rules of the code from the signs on the screen, and at utilizing the possibilities of manipulation that they are offered: 'cheats', 'walkthrus', and editors to enhance the possibilities of their avatars are important instruments to improve the game. While Umberto Eco's concept of the open text is probably old news to gamers, literary scholars seem to forget the achievements of literary theory in overcoming the notion of an autonomous text when it comes to applying this concept to computer games.

Thus, an approach concerned only with the signs generated on the user interface fails to extract all possible meanings from the text. In my opinion, this approach misconstrues the reading direction in computer games: while traditional narratives tend to make the reader forget that he is reading a text through which the experience of a fictional world is mediated, computer games constantly supply the player with references to the grammar that governs the seemingly unmediated experience of the makebelieve world she is immersed in. Meta-fictional signals are also to be found in printed literary texts, of course, but there is one important difference: in traditional narratives these signals will not make the text appear as something that can be manipulated by the reader, while computer games challenge the player to a subversive reading strategy. When scholars of literary studies disregard these possibilities for manipulation, they retreat to the position of a 'naïve' reader who believes in the author's omnipotence in determining the narrative, and for whom only a typological exegesis makes sense. This 'paranoid' approach has long been transcended by the players of computer games, who acknowledge the arbitrary manifestations of the code as something that can be manipulated at will.

3. GAME AND READING

Yet the approach of literary studies is not only in danger of confusing text and reading, but also in regarding text and narrative as equal. After all, it is nearly impossible to make the narratives of computer games fit the Aristotelian definition of something with a beginning, a middle, and an end. And although literary studies are concerned with non-narrative texts as well, it is a typical mistake to describe computer games in a fashion that allows them to fit this schema. While Jørgen Kirksæther must be credited with trying to differentiate between game and reading, he too becomes caught up in the Aristotelian definition of narrative. By postulating that "[the] middle is the really interesting part," he is led to conclude that this middle is a sequence of circular internal narratives that are embedded in the frame of a reading and that it is inevitable to play certain passages of a game over and over again, until the player is granted access to the next level. His acceptance of the code's absolute authority necessitates the conclusion that "[t]he appeal of games isn't in mastering a complicated set of controls, but rather in submitting to a set of rules and trying to accomplish something under these rules' restrictions." [4] Kirksæther concedes that he regards it as impossible to separate the graphic interface from the logic and the

structure of the game. But thus he renders the differentiation between text and reading futile.

As we have seen, literary studies broaching the topic of computer games as if they were interactive narratives or interactive films tend to adjust their object of analysis according to the means they have at hand. But if their analysis is to produce valid results, scholars approaching the subject of computer games from a literary perspective must be aware of this tendency and they must ensure that the instruments employed are appropriate to this subject. Espen Aarseth must be credited with outlining the dangers of exporting the terminology of literary studies to a new field of study:

[...] I wish to challenge the recurrent practice of applying the theories of literary criticism to a new empirical field, seemingly without any reassessment of the terms and concepts involved. This lack of self-reflection places the research in direct danger of turning the vocabulary into a set of unfocused metaphors, rendered useless by a translation that is not perceived as such by its very translators. [...] Even if important insights can be gained from the study of extraliterary phenomena with the instruments of literary theory (cautiously used), it does not follow that these phenomena are literature and should be judged with literary criteria or that the field of literature should be expanded to include them. In my view, there is nothing to be gained from this sort of theoretical imperialism, but much to lose [...] [5]

That is why, in assessing the possibilities of literary terminology for analyzing computer games, I will also try to show the limitations of this approach. Furthermore, I will highlight the aspects of this field for which there is no appropriate terminology as of yet. For the development of new critical terminology I depend on the work of other scholars in this field, for it is only through a critical assessment of their accomplishments that this project can be realized.

4. AESTHETIC CRITERIA

To date there have been few approaches to the field of computer games from a broad cultural perspective. The establishment of institutions such as the Computerspielemuseum in Berlin, and events such as *LaraCroftism* in Munich, or the *Computer Games and Digital Textualities* conference in Copenhagen, show that there is a growing number of scholars working in this field. Furthermore, the publication of books such as David S. Bennahum's *Extra Life* and Steven Poole's *Trigger Happy* make it obvious that computer games have become a part of our culture and that there is public interest in discussing this phenomenon in a broader context.

Therefore, a serious approach to computer games which attempts to develop an appropriate terminology should be concerned with developing independent aesthetic criteria; i.e. independent of the criteria established in commercial reviews and independent of the criteria employed in the criticism of other media. As early as 1993 Ted Friedman pointed out the need to develop a 'software theory', and he stressed the role of computer games in this enterprise: "One area that has received scant attention from cultural theorists, however, offers particularly fertile ground for inquiry: the world of computer games." [6] Friedman is hardly the first to draw attention to this area – in 1985 Mary Ann Buckles wrote her doctor's thesis on the game *Adventure* [7] – but he must still be

credited with pointing out the limitations of a literary approach to this field. Consequently, Friedman concentrates not on a narrative genre of computer games, but rather on strategy games such as SimCity, marketed not as a computer game but a 'software toy' by its manufacturer. Although Friedman is aware that SimCity has been derived from the level generator of another computer game, he differentiates between the possibilities for manipulation offered explicitly by the software, and those that might be regarded as 'inofficial' strategies such as 'cheats'. Mostly due to this discrimination, Friedman agrees to Orson Card's claim that "[e]very freedom you can give to the player is an artistic victory."[8] Establishing a poetics of computer games, rather than an aesthetics, cannot be the aim of a critical 'software theory', however. Games in which the player is required follow certain guidelines cannot be excluded from analysis because they do not offer as many possibilities for interaction at first glance. Notwithstanding the establishment of specialized areas within this field, a theory of computer games should initially keep its focus as wide as possible. Such an approach has been suggested by Espen Aarseth. In his book Cybertext he defines the term 'cybertext' as follows:

The concept of cybertext focuses on the mechanical organization of the text, by positing the intricacies of the medium as an integral part of the literary exchange. However, it also centers attention on the consumer, or user, of the text, as a more integrated figure than even reader-response theorists would claim. The performance of their reader takes place all in his head, while the user of cybertext also performs in an extranoematic sense. During the cybertextual process, the user will have effectuated a semiotic sequence, and this selective movement is a work of physical construction that the various concepts of 'reading' do not account for.[9]

As Aarseth points out, the characteristic feature of cybertexts is that they are 'ergodic', a term borrowed from physics and put together from the Greek words for work and path: "In ergodic literature, nontrivial effort is required to traverse the text." Obviously, this term includes not only computer games, but also hypertexts, MUDs and MOOs, as well as a number of printed texts ranging from the ancient Chinese I Ching to Raymond Queneaus's Cent mille milliards de poèmes. In his essay "Aporia and Epiphany in Doom and The Speaking Clock" Aarseth stresses the importance of an inclusive definition: "The worst kind of mistake an aesthetic theory of ergodic art can make is to assume that there is only one type with which to be concerned [...] with a single set of properties."[10] But for a study exclusively concerned with computer games this definition has to be modified. Aarseth's definition of ergodic texts combined with Friedman's insistence on regarding software in its own right, i.e. as a unique form of aesthetic expression, supplies us with a working definition of what a computer game is. Thus, games with a graphic interface are included as well as 'text adventures', while interactivity might refer to the game as well as the code. Nevertheless, this definition might require further modification in the future.

5. COMPUTER GAME GENRES

A first attempt to establish independent aesthetic criteria can be made by regarding computer game genres. The differentiation we find in popular computer game discourse seems rather arbitrary, but since any other way to classify them would be equally arbitrary, I think it makes sense to analyze the existing genres rather than create new ones. There are five basic genres taken into account: action games, adventure games, role playing games, simulation games, and strategy games. A closer look at these genres reveals that they can be differentiated by the following three criteria: narrativity, openness, and interactivity. The genres can be placed in a triangular matrix that is defined by these criteria, since they appear to be complementary (FIGURE 1) Thus, adventure games, for example, rank high in narrativity, but by the same token their openness and interactivity are reduced by the game's inherent narrative structure. Strategy games, on the other hand, have a high level of openness, for there is usually no order in which to complete certain tasks, and the possibilities for interaction are often numerous. The result of this open structure is that there is no real narrative to speak of, and the frequency of interaction is relatively low. Contrarily, action games usually have a very high frequency of interaction, while the range of these actions is rather narrow. Thus, action games are characterized by a high level of interactivity, and a low level of both narrativity and openness. Role playing games and simulation games take intermediate positions in the matrix.



One way to judge a game by aesthetic criteria is to see how well it succeeds in fulfilling the criteria of its genre. While this is certainly not all there is to it – often an aesthetic success depends on breaking the rules rather than submitting to them - the criteria that define the different genres deserve more attention. First of all, we have to differentiate between openness and interactivity, since they are easily confused. Both criteria refer to the 'freedom' of the player in the fictional world of a computer game. But while interactivity refers to the frequency of the player's interactions with this world, openness refers to the range of different interactions offered by the game. It is important to keep in mind, however, that both criteria apply to the interface as well as the underlying code. Since games that offer a high level of openness at the interface usually allow quite a bit of manipulation on the level of the code as well, it is futile to differentiate between these levels at this point. But the interplay between code and interface highlights another important aspect. Since it is possible to increase the level of interactivity by exhausting the full spectrum of interactive possibilities the game offers, it is obvious that these criteria are interdependent. The level of interactivity and openness might even change during the game - but an increase in openness will necessitate a decrease in interactivity and vice versa. Therefore, the borders between the genres are fluent, and especially hybrid genres are often difficult to categorize.

6. GAMES AND NARRATIVES

From a literary perspective, narrativity is the most interesting and the most problematic of the three criteria. A narrative-oriented game is not more prone to be analyzed from a literary perspective, despite the literary scholar's preference for games of the more narrative genres. After all, the field of literary studies extends beyond the narrative genres of literature and includes dramatic and poetic texts as well as scientific and philosophical texts. Yet the limitations of the philological approach to computer games become very obvious when it comes to making sense of non-narrative games. While it might still make sense to compare adventure games with medieval quest narratives, or action games to certain epic genres, it would be hard to argue that *Tetris* is an interactive poem. While it should certainly not be the aim of literary studies to assimilate computer games in such a way, we must keep in mind that by simply stating that a narrative element exists within them, these comparisons are implicitly made.

So, what terminology does literary theory supply us with to describe narrativity in computer games? Obviously, most games tell a story, even if it can often be summarized in a couple of words. Even non-narrative computer games are often set within a narrative, paratextual frame. Jørgen Kirksæther quotes an example from the instructions for the game *Silkworm*: "[...] Earth's chances of survival hang on [sic] a thread, a thread so gossamer fine that it could be made of silk. Realizing this, the weapons scientists codenamed civilisation's last stand Operation Silkworm. Step forward hero, read the briefing and take the controls...," and he continues: "Now, if we for a minute can put aside the rather ridiculous explanation of the game's name, what happens here? I'd say three things: 1) You're being told the beginning of a story, 2) you're being invited to actively take part in it, and 3), it's quite obvious that the story isn't over."[11]

The implications of this statement are somewhat peculiar when computer games are regarded as hypertexts, as by George Landow, or as interactive films, as in Kirksæther's approach. While Landow is lead to conclude that this "calls into question ideas of plot and story current since Aristotle" [12], Kirksæther encounters serious difficulties in keeping up the Aristotelian definition of narrative, resulting in contradictions he cannot resolve. Neither approach produces a convincing definition of narrativity in computer games. When we turn to Espen Aarseth's concept of cybertext once more, we find that he is the only one who regards "[t]he adventure game [as] an artistic genre of its own, a unique aesthetic field of possibilities, which must be judged on its own terms." Aarseth also seems to support the thesis that the reading direction is reversed in computer games when he claims: "In the determinate cybertext [...] the functions of plot (szujet) and story (fabula) appear to have traded places, somehow." Yet he continues:

But this is not exactly the case. The concept of plot is unsettled by the reader (user), who, being strategically within it, is in no position to see through it and glimpse a story behind. It is often argued that narrative plot is also something that is only discovered or reconstructed by the reader after the end is reached; and this could be seen to imply, contradictory to my argument, that there is no great difference between the narrative and the ergodic situation as far as plot is concerned. But there is a difference, and for a very simple reason: the bewildered reader of a narrative can safely assume that the events that are already encountered, however mystifying, will make sense in the end (if the plot is to make sense at all); whereas the player of an adventure game [...] is not guaranteed that the events thus far are at all relevant to the solution of the game.[13]

This claim is easily contradicted by pointing out that printed literary texts tend to supply their readers with irrelevant information as well. It might just as well be argued, however, that irrelevant narrative elements are clearly considered an aberration in narrative theory, while it is quite usual for the player of a computer game to be confronted with seemingly irrelevant information. In some games, this strategy of misinformation may even be regarded as one of the principal structural elements.

Aarseth's conclusions from this thesis are rather dramatic: "[The adventure game] effectually disintegrates any notion of story by forcing the reader's attention on the elusive 'plot'. Instead of a narrated plot, cybertext produces a sequence of oscillating activities effectuated (but certainly not controlled) by the reader.' [14] He suggests to replace the term 'story' with 'ergodic intrigue', signifying an element that structures and controls the adventure game. The ergodic intrigue is directed against the 'intriguee', a role that Aarseth equals with the implied reader and the protagonist of printed narrative texts. This is consistent Aarseth's argument that the roles of the protagonist and the implied reader converge in adventure games, an argument also put forth in Brenda Laurel's [15] and Marie-Laure Ryan's [16] work. It is difficult to follow him, though, when he concludes that "[t]hus, the determinate cybertext reconfigures literary experience along a different plane than the narrative. Instead of a narrative constituted of a story or plot, we get an intrigue-oriented ergodic log - or to adopt Gérard Genette's and Seymour Chatman's term, ergodic discourse." [17] The problem of narrativity in computer games is merely shifted to another level in this model - since Chatman proposes that signs play the role of a mediator between story and plot. Therefore, Aarseth is unable to differentiate between narrative and ergodic discourse without introducing two new narrative concepts that he calls the 'event plane' and the 'progression plane':

In a narrative, the discourse consists of the event plane, where the narration of events takes place, and also what I call the progression plane, which is the unfolding of events as they are received by an implied reader. [...] In adventure games, the relation between events and progression is defined by a third plane of discourse: a negotiation plane, where the intriguee confronts the intrigue to achieve a desirable unfolding of events. [18]

So, after plot and story have been dismissed as irrelevant for the analysis of computer games, they are reintroduced as newly defined planes of discourse. The model that Aarseth presents us with actually differs from Chatman's model in only one detail. The mediating plane that is called 'discourse' in Chatman's, and 'negotiation plane' in Aarseth's model, is interactive. Unfortunately, this model does not supply us with an explanation of the connection between interactivity and narrativity either, for they are regarded as inseparable parts of the negotiation plane. A more promising model is supplied by Janet Murray, who differentiates between traditional printed texts and interactive texts by regarding them under the aspects of 'agency', 'rapture' and 'immersion':

These 'three key pleasures' [...] are uniquely intensified in electronic media. [...] Immersion [...] is 'the sense of being transported to another reality, such as a game world. Rapture is the 'entranced attachment to the objects in that reality' – in other words, the addictive trance that gamers fall into for hours at a time. And agency is 'the player's delight in having an effect on the electronic world,' which is possible because the player is a free agent who can make choices. [19]

7 AGENCY, IMMERSION, AND RAPTURE

The term agency must be regarded as more than just an alternative term for interactivity, since it encompasses elements of narrativity and openness as well. Therefore, Murray's terminology seems to be predestined to take a closer look at the connections between these three categories. Being more than just a duplication of our original triad, we must assume a more complicated relation between the two sets of categories. Indeed, it makes much more sense to put these categories into a causal relation (FIGURE 2): Thus, rapture results from the combination of narrativity and openness, while immersion stems from combining interactivity with narrativity. Agency, on the other hand, is the result of the combination of interactivity and openness. Narrativity and agency are opposed to each other, since there is hardly a way to regulate a narrative in a game that grants the player great influence on the course of his actions. Immersion is opposed to openness, because a great range of possible actions tends to make a game more abstract, and thus anti-immersive. And rapture is opposed to interactivity, since a high level of interactivity will effectively counteract ,,the entranced attachment to objects in that reality."



FIGURE 2

Narrativity is thus connected to openness and interactivity, respectively, by a relation that results in rapture on one hand, and immersion on the other. Murray's observation that the traditional 'pleasures' of the text are intensified in electronic media is consistent with this model, since in interactive texts such as computer games these pleasures are enhanced by openness and interactivity. In this light, we must reconsider Aarseth's thesis that plot and story are of little relevance to computer games, whereas the importance of discourse is increased in an interactive environment. It is clear from the above model that rapture and immersion decrease the reader's critical distance towards the text, thus counteracting her ability to discern the unfolding of events (story) and the temporal and causal structure (plot). The same tendency can be discerned in printed text, even though it is not quite as strong. But if rapture and immersion are responsible for plot and story being of little relevance, they must also be responsible for stressing the discourse in computer games. And indeed, games in which there is an equilibrium of openness and interactivity, thus offering a maximum of agency, tend to disregard the underlying narrative structure in favor of the momentary events and the way these events are related. That this description characterizes simulation games very accurately is consistent with Marie Laure Ryan's statement that a simulation

system is "not a narrative, but a narrative matrix" [20]. The level of narrativity is necessarily very low in a narrative matrix, for in such a structure there is no internal hierarchy that necessitates a certain narrative direction.

8. PERSPECTIVE

It is impossible to analyze the narratological strategies in computer games without regard to the perspective from which the events of these games are being related. In other words: we have yet to answer the questions of point of view. This question is intertwined with two highly problematic questions in respect to narrative control, i.e. what it is exerted by, and whom it is exerted on. But before an attempt is made to answer these questions, we have to discuss how the terminology of literary theory can contribute to analyze narrative roles and point of view. At first glance it seems simple enough to rely on a traditional approach such as Stanzel's classical model of narrative situations. After all, terms such as 'first-person shooter' and 'third-person shooter' seem to be derived from terms coined in narrative theory. These models do not account for the difference between the narrator and the observer, though. Because a differentiated analysis depends on a differentiated model of point of view, Gérard Genette's work should be taken into account, since he was the first to point out that "most studies of point of view [...] treat two related but different questions as if they were interchangeable. [...] [T]hese questions are 'who sees?' v. 'who speaks?'" [21]

In trying to answer the first question – who sees? – we have to understand through whose eyes the player of a computer game perceives the fictional world presented to her on the screen. It seems all too easy to answer this question in respect to games in which the player sees quite literally through the eyes of a character in this fictional world, as it is the case in 'first-person shooters', such as *Doom* or *Quake*. In these games the player is represented through an 'avatar', similar to the way a 'first person character focalizer from within' represents the reader in a traditional narrative. In this terminology, 'first person' refers to the perspective from which the fictional events are narrated, while 'character focalizer' means that the narrator is actually involved in this narrative. A narrative related 'from within' will not grant the reader access to information that is not known to the narrator. This perspective has certain characteristic limitations: a limited overview of the temporal and spatial dimensions of the fictional world, and limited knowledge about what is going on in this world. Yet there is something of an objective normative system in the form of implicit rules that define what is 'right' and what is 'wrong' within the limits of the fictional world. This draws our attention to the fact that seeing often implies being seen, and this raises the question of in which way this view on the player is implemented in the game. Apparently, this manifestation is in some way connected to the code of the game, and thus we are tempted to identify it with an 'implied author'. But the implications of this assumption are problematic, since the role of the implied author is at least partially transferred to the player in a computer game. The ensuing dilemma can be resolved in two ways. Either the role of the implied author is divided into two different roles, one of which is taken by the player, while the other is part of the code, or the division between (implied) author and (implied) reader is regarded as non-existent as in Kirksæther's model.

Yet the problem of mediation can be avoided in its entirety if we regard computer games as an immediate form of communication.

Brenda Laurel suggests to regard computer mediated communication as a form of dramatic interaction. In this model, instead of a convergence of the roles of reader and author, the communication process is conceptualized as a convergence of spectator and actor. Thus, the player has some 'creative freedom' in fulfilling his role, although he is limited by the possibilities offered to him by the text, or code: "The users of such a system are like audience members who can march up onto the stage and become various characters, altering the action by what they say and do in their roles." [22] This model could also be adapted for computer games in which the player does not perceive the gameworld through the eyes of his avatar, but in which he actually sees a representation of his character that he controls in a fashion similar to directing a play. While it seems simple enough to transpose this model into narratological terminology by transferring the role of focalization to the player, the convergence of the narratological and the dramatic model becomes more complicated when we turn to the question 'who speaks?'

After all, Laurel's model does not include a narrator, and even if we assume a 'dramatic narrator', this poses the same problems that we tried to avoid by turning to the concept of *Computers as Theatre*. The implications of consolidating the narratological and the dramatic models become even more puzzling when turning to games in which there is simply no way of identifying the player with one of the *dramatis personae*, games such as *Popolous* or *SimCity* in which the events are being focalized externally. In these games, the only 'voice' we can identify is the player's voice, which is giving instructions to the inhabitants of the fictional world. If the player identifies with anything at all, it is a whole tribe, people, or nation. This special problem might be resolved, however, by a suggestion brought forth by Ted Friedman in "Making Sense of Software":

We could see playing SimCity, then, as a constant shifting of identifications, depending on whether you're buying land, organizing the police force, paving the roads, or whatever. This, I think, is part of what's going on. But this model suggests a level of disjunction – jumping back and forth from one role to the next – belied by the smooth, almost trance-like state of the gameplay. Overarching these functional shifts, I think, is a more general state of identification: with the city as a whole, as a single system. [23]

Therefore, in order to answer the question of what role the player takes in the communication process of a computer game, we have to take a closer look at what role identification plays in this process. And it is only in relation to the position of the player in the communication process that we can determine the position of the 'narrator'. But this can only be accomplished if we keep in mind that the narrative equivalent of the point of view in a 'firstperson shooter' is actually a narrative in which the reader is addressed in the second person singular. The text-based adventure game Zork, for example, begins as follows: "You are standing in an open field west of a white house, with a boarded front door." It seems paradoxical to equate a point of view from which the player sees through the eyes of an avatar with a narrative perspective that directly addresses the player. Yet this apparent paradox is the key to understanding the basic communication process of computer games, because it is obviously not the player herself who is addressed in such a way, but the 'narratee'. This element of the communication process is not identical to the implied reader, although Aarseth seems to suggest such an equation when he identifies the 'intriguee' in adventure games as the implied reader.

Yet the way in which the narratee is addressed is clearly not metafictional, as opposed to the meta-fictional strategy of directly addressing the reader in traditional narratives. On the contrary, this means is employed to simplify the 'willing suspension of disbelief', i.e. the player's immersion. The narratee, on the other hand, can be equated easily with the spectators of a play, since the actors on stage can address the audience without leaving the fictional frame by pretending the spectators are part of the fictional world.

9. COMMUNICATION

Apparently Laurel's model can be reconciled with a modified model of communication in narrative texts after all. Our original thesis - that the reading direction in computer games is reversed can then be abandoned as trivial. The unified communication model does not supply us with the means to determine the position of the 'narrator'. This can only be accomplished by taking a look at Aarseth's model of a computer game's principal components (FIGURE 3). In this model, the game's code is mediated through an interface that fulfills the two functions of analysis and synthesis. It is through this interface that the player is able to communicate with the code at all. The communication between the code and the interface is further mediated through a 'simulation engine' and a 'representation engine'. These engines are the parts that have to be taken into account when we want to locate the 'narrator' in this communication process, because they are the parts of the code that can be manipulated by the player, whether directly or indirectly. It is worth noting that this model is dialogic, which makes it seem worthwhile to approach the field of computer games with the terminology developed by Mikhail Bakhtin, as suggested by Geoffrey Rockwell.





Before we turn to the implications of this approach, though, we have to consider the role that the representation engine and the simulation engine play in the player's dialogue with the code. It is clear from Aarseth's model that representation corresponds to the 'voice' addressing the player in text-based adventure games. This voice is mediated through the interface that synthesizes the output of the code in order to make them 'legible' for the player. The precondition for the player's ability to make sense of these data is the player's identification with his avatar, or narratee. This is necessary, because the signs produced by the interface only make sense in the context of the fictional world. With this precondition fulfilled, the player can react to these signs appropriately and manipulate the interface in such a way that that makes sense to the analytic device of the interface. This input is then passed on to the simulation engine. Thus, the representation engine is confronted with a narratee that is not only 'listening', but answering as well. This part of the narration process is not at all 'auctorial', but rather equal to the narratee in respect to it's information about the game-world, because it is only granted access to the data that are being transferred to it from the game's code – either directly, or mediated through the simulation engine. Therefore, we can equate the code with an implied author that determines which information is being made available to the narrator. We must then assume two narrative voices, one of which is represented by the avatar – a 'character focalizer from within' – while the other is a 'narrator focalizer from without' that can not be located within the communication process, but must instead be considered an effect of the interplay of the different narrative elements.

The internal normative system, on the other hand, that reinforces 'good' behavior, and punishes 'bad' behavior, is to be located within the game's code itself, using the representation engine merely as an executive organ of its measures. But due to the dialogic structure of the communication process, it is up to the player if he subscribes to these values or not. A drastic example for such denial of a game's internal values are those players of Ultima Online who no longer strive for wealth, adventure, or social status, but rather spend their time killing other player's characters. Contrarily, many Quake-clans subscribe to even stricter rules than those supplied by the game's code. Thus, ambushing an opponent from a secluded spot (camping) is regarded as dishonorable, even though the game's code reinforces such behavior. So the multiple voices within the game can disagree, or even contradict each other. Considering these multiple voices, it seems worthwhile to take a closer look at Rockwell's approach to questions of identification and control with Bakhtin's terminology.

10. TAKING CONTROL

This approach seems especially promising in regard to two questions that Rockwell addresses: Firstly, "[what] types of characters that interact in the game with special attention to the character the player is allowed to develop," and secondly, "the types of interactions that can be performed in the game with special attention to the interactive possibilities for the player." [24] Rockwell points out that the question of interaction is closely related to the question "How are you defined by the choices the game affords?" This question implicitly assumes the equality of identification and interaction. This seems to imply that the player cannot fully identify with his avatar unless he ackowledges this character's limitations. Is identification achieved through interaction, then? Yes and no. On one hand, immersion is undeniably increased by the player's opportunities to act upon the game-world, and identification through action has great suggestive power. On the other hand, the player's realization of his limited possibilities necessarily counteracts the player's immersion - thus weakening the player's identification with the narratee. Therefore, the fact that the player's avatar is unable to do certain things appears to be a hindrance for identification with this character. Consequently, Friedman describes the process of playing a computer game as a process of demystification: "Learning and winning [...]a computer game is a process of demystification: one succeeds by discovering how the software is put together." [25]

Thus, 'solving' a computer game appears to be a process in which the player learns to decipher the signs on the interface as manifestations of the rules as determined by game's code - in a way, this can be regarded as a reversal of the reading direction after all. It is only through this process of demystification that the player can gain control over the game – subversive reading strategies such as 'cheats' included. The process of analyzing the inherent rules of the code that constitutes the game can then be regarded as a hermeneutic process. Only in becoming aware of the full extent of his possibilities can the player master the game – simultaneously breaking the spell of the game through the sacrifice of its immersive power. This must not be regarded as a disadvantage: the player can still recreate the magic of the game by taking refuge to the willing suspension of disbelief.

REFERENCES

[1] MONTFORT, NICHOLAS A. (1995): "Interfacing with Computer Narratives – Literary Possibilities for Interactive Fiction" [http://nickm.www.media.mit.edu/people/nickm/srthesis]

[2] ANDERSEN, PETER BØGH (1990): A Theory of Computer Semiotics: Semiotic Approaches to Construction and Assessment of Computer Systems, Cambridge University Press, Cambridge

[3] DALUM, ASTRID and SØRENSEN, FINN V. (1996): *Interactive Fiction* [http://www.centrum.dk/users/finnv/intfict.htm]

[4] KIRKSÆTHER, JØRGEN (1998): "The Structure of Video Game Narration" [http://cmc.uib.no/dac98/papers/kirksaether.html]

[5] AARSETH, ESPEN (1997): *Cybertext - Perspectives on Ergodic Literature*, The Johns Hopkins University Press, Baltimore and London

[6] FRIEDMAN, TED (1993): Making Sense of Software: "Computer Games and Interactive Textuality" [http://www.duke.edu/~tlove/simcity.htm]

[7] BUCKLES, MARY ANN (1985): Interactive Fiction as Literature: The Storygame 'Adventure', Ph.D. diss., University of California, San Diego

[8] CARD, ORSON SCOTT (1989): "Gameplay", in: Compute, April 1989

[9] AARSETH 1997

[10] AARSETH, ESPEN (1999): "Aporia and Epiphany in *Doom* and *The Speaking Clock*: The Temporality of Ergodic Art", in: Marie-Laure Ryan (ed.): Cyberspace Textuality, Indiana University Press, Bloomington and Indianapolis

[11] KIRKSÆTHER, JØRGEN 1998

[12] LANDOW, GEORGE (1997): *Hypertext 2.0. The Convergence of Contemporary Critical Theory and Technology*, Johns Hopkins University Press, Baltimore and London

[13] AARSETH 1997

[14] ibid.

[15] LAUREL, BRENDA (1991): Computers as Theatre, Addison-Weasley, Menlo Park, CA

[16] RYAN, MARIE-LAURE (1994): "Immersion versus Interactivity: Virtual Reality and Literary Theory", in: *Postmodern Culture* [http://calliope.jhu.edu/journals/pmc/] ------ (1997): "Interactive Drama: Narrativity in a Highly Interactive Environment", in: *Modern Fiction Studies* 43/3

[17] AARSETH 1997

[18] ibid.

[19] MURRAY, JANET (1997): *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, The Free Press, New York

[20] RYAN 1994

[21] RIMMON-KENAN, SHLOMITH: *Narrative Fiction. Contemporary Poetics*, London 1983

[22] LAUREL 1991

[23] FRIEDMAN 1993

[24] ROCKWELL, GEOFFREY (1999): "Gore Galore: Literary Theory and Computer Games" [http://www.interchange.ubc.ca/winder/papers1999.htm]

[25] ibid.

Designing Interactive Narratives

Mikael B. Skov Department of Computer Science Aalborg University DK-Fredrik Bajers Vej 7, 9220 Aalborg East +45 +9635 8080

dubois@cs.auc.dk

ABSTRACT

Aspects of interaction and narration challenge the design of interactive narratives. Research has indicated that the design of interactive narratives is a difficult and challenging task. Interactive narratives exhibit features and characteristics from story telling or film and differ from more conventional software systems by being functionally more complex and by addressing a broader and wider audience of users. Interactive narratives include software systems like computer games or training and assessment systems. This paper discusses the inherent contradiction between interactivity and narrativity. We exemplify solutions of integrating interaction and narration in computer games and we illustrate the potential problems associated these structures. Finally, we propose an additional approach to interactive narratives design by creating worlds to navigate in instead of designing strict narrative structures.

General Terms

Design, Experimentation, Human Factors.

Keywords

Interactive narratives, design processes, narrative structures, computer games

1. INTRODUCTION

The blooming genre of computer games is continuously becoming more and more advanced and their role in the software industry is becoming more and more dominant. The computer game market revenue is expected to increase to a total of 6.54 billions of dollars by the year of 2002, cf. [14]. This is equivalent to an increase of more than 36% from the total revenue in 1996. Computer games are often interesting to study since people learn to operate them easily without training and often develop strategies for improving their performances with the game, cf. [15]. Computer games can be very different with respect to content, purposes, or user groups, however many computer games contain a great deal of narrativity or story telling. Kolstrup describes and analyses three computer games, *Crosstown, Tomb Raider*, and *Riven*, as narratives and states that these three computer games are much closer to

First published at CO SIGN-2001, 10-12 September 2001, CWI, The Netherlands. Peter Bøgh Andersen Department of Computer Science Aalborg University DK-Fredrik Bajers Vej 7, 9220 Aalborg East +45 +9635 8080

pba@cs.auc.dk

traditional narratives than e.g. newspapers and television [20]. Also, Konzack analyzes the *Barbie Fashion Designer* game as an interactive narrative [21]. Since the release of the computer game *Half-Life*, the value of story telling or narration in computer games has been convincing, cf. [22]. According to Rollings and Morris, *Half-Life* does not tell you a story but enables you to create one yourself.

The design and creation of interactive narratives is often complex and difficult, cf. [12,13,14,28]. Most contemporary development methods for analysis and design of computer-based information systems are intended for either administrative or technical embedded software, cf. [2,18]. But few interactive narratives can be regarded as administrative or technical software, cf. [28], they belong to completely different genres. Also, computer professionals, e.g. designers, programmers, database experts, human-computer interaction experts, have primarily conducted the design of conventional information systems or interactive systems in collaboration with work domain people. However, technologies such as computer games or other kinds of new media do not necessarily have a corresponding work domain associated and hence these technologies pose new requirements on the design process and puts new problems on the agenda.

Game designer Chris Crawford aptly identifies one such problem, the inherent conflict between narration and interaction:

... the basic conflict emerges because the artist insists on taking the audience down a predetermined path (as is the case with conventional stories), while at the same time demanding the audience's active involvement in the course of the experience. The solution is for the author to relinquish control of the path to the audience... Crawford [8]

His solution may be the right one, but how can it be approached and designed? Many techniques for engaging the consumer of a media product presuppose that the author has full control over the product; if this presupposition no longer holds, the techniques are no longer available, and the product easily becomes boring and uninteresting. Take movies as an example: According to Bordwell, the film-viewer continually creates expectations about what will happen or has happened [3]. These expectations has the form of schemata (a murder is possible), that generates hypotheses about what one should look for next (vulnerability of the victim); the hypotheses guide the interpretation of what is actually perceived (a knife is interpreted as a weapon, not as a kitchen utensil), and cues in the material fill out slots in already existing schemata (Mr. Schmidt turns out to be the murderer), or generate new ones (not a crime, after all, but a joke). If such expectations are not created, there will be no suspense or curiosity, which again means that there will be no engagement. The problem is not to create expectations (*set ups*) but to ensure that expectations are fulfilled — or explicitly disappointed in order to create surprise (*pay offs*), since, if the author does not know which path a user will take, then how can he plan payoffs?

Another basic technique is to hide information to the audience in order to create curiosity and then disclose it at a later time. The whole genre of detective novels is dependent upon this technique. However, how can the author be sure that the user will not see the information at time T_1 but that he will encounter it at time T_2 , given that it is the user who is in control of the system?

In his specification of the *Erasmatron* game engine, Crawford has in fact provided for this technique:

...most of the time you don't care about the audience, but some verbs won't take place in front of the wrong onlookers. For example, you don't knife DirObject in the back in front of a crowd; and you don't punch out that bully unless your girlfriend is watching so she can be impressed. The story engine respects whatever audience requirements you have specified for the verb in the General display... Crawford [9]

Thus, the author not only needs control over "story time" (the events in the fictive world) but also over "plot time" (the sequence in which the events are presented to the audience). There are also global patterns of action that contribute to our sense of pleasure and fulfillment. The popular curve of tension shown in figure 1 underlies millions of narratives, ranging from Greek tragedy to modern soap operas.



Figure 1. The Aristotelian curve of tension.

Some theoreticians have suggested that we use the opportunity to throw away all these hoary heirlooms, and explore the new medium with fresh eyes: the postmodern, nonlinear open work. This is probably a fruitful approach for the experimenting avantgarde arts, but it may not be very successful if we are concerned with the medium's potentiality as a mass medium, where computer games for many years have set the agenda.

The next section describes techniques and narrative structures that are actually used for handling these difficulties. They are ordered according to their degree of interactivity, the lowest ones first.

2. DESIGN EXPERIENCES

As stated above the inherent problem of combining interaction and narration imposes potential problems on the design of interactive narratives. Design experiences from two interactive narrative authors will serve to illustrate how the design process can be conducted for interactive narratives and how different narratives structure can be applied.

The two authors were interviewed and design documents were used as illustration of design approaches. The interviews tried to address their practical experiences in design processes of interactive narratives. Both interviews were planned and conducted as semi-structured interviews. The primary focus during the interview was on activities, concepts, and products applied during a typical design process. The interviews had a duration of approximately one hour each and were taped on a Dictaphone. Afterwards both interviews were transcribed word for word.

The first author (from now on called the game author) has been involved in several major computer games productions and of productions of other story-telling applications. He is associated to a number of Danish software houses specialized in developing computer games both for children and adults, e.g. Savannah A/S and Deadline A/S^1 . He has worked with interactive narrative design for more than five years and have been involved in many design projects. He holds a masters degree in communication. Also, he has worked with other aspects of narrativity, e.g. published a collection of poems and worked with a major Danish television broadcasting company on an interactive narrative project. The game author will primarily address two kinds of computer games in the following sections. First, he has worked on productions of computer games for children. These games can be characterized as being fantasy worlds where a number of characters live and behave. The user of these games (usually a child) can act as more figures in the world where each figure is associated a number of behaviors that has to be applied in order to solve the problem of the game. Usually, the child has to help find something, e.g. a missing friend. The child can interact with other figures in the world and can move between different locations in order to solve the problem, e.g. move from the living room to the basement. The second class of games referred to by the game author is primarily aimed at adults. These are more detective games where the user has to solve some sort of a puzzle. Again, the user is able to move between different locations, this time in a city where each location will provide the user with some information on the puzzle. Both classes of games rely on the fact that user can explore the worlds they are placed in thereby discovering new things themselves and also in this sense create their own stories.

The second author (from now on called the multimedia author) works in a Danish software house called Interact A/S², specialized in designing and producing interactive narrative systems for training and assessment of work situations. He holds a master's degree in English and has worked with creation of interactive narratives for more than four years. Additionally, he has worked with theater and other story-telling performances for more than 8 years and he has published a short story as a textbook. The interactive narratives referred to by the multimedia author are primarily aimed at assessing or training work domain people. E.g. he has authored a training system for sales people in a major shop in Denmark. The purpose of these systems is that the user should

¹ See (http://www.savannah.dk/) and (http://www.deadline.dk/)

² See (http://www.interact.dk/)

learn how to handle various situations from typical sale situations in the shop. In this sense, the sales person should learn what to do and what not to do. Usually, the systems contain a number of video sequences that can be played to the user showing a particular situation. After each video sequence the user is required to select a solution to the problem depictured by the video sequence. In the end, the user will be able to identify in which situations he or she did the wrong thing and hopefully learns from that. The evaluation is often conducted with help from other people, e.g. people from the management or a consultant.

2.1 Narrative Creation

The fundamental task during the narrative creation is to conceptualize the overall idea of the future interactive narrative. The purpose is to establish a common ground for the further work in the design projects and set up specific requirements directed towards a customer or towards the rest of the design team. The multimedia author tried to address the key concept for the future training and assessment system by formulating key requirements to the system. This would usually be done based on a brainstorm of ideas between him, a producer, and the company.

... based on the brainstorm I try to require as much information as possible in order to design a concept for the narration ... I would normally not write anything down from this meeting for myself, it is all about listening and understanding ... Multimedia Author

He would also try to determine at an abstract level the underlying narrative structure for the system (narrative structures are explained and illustrated in the next section). Finally, he would advice the programming team on the concept for the narration, but normally not initiate any direct programming tasks. The game author, on the other hand, tries to conceptualize the overall idea of the future computer game by creating a document written in a rich and colorful language. This is primarily done in an attempt to illustrate and define the atmosphere of the game. He is often faced with a different situation since he has to come up with a fictitious world.

When the future system has been conceptualized a deeper understanding and description of the situation the future interactive narrative has to be provided. The primary outcome could be a short story telling the narration at a somewhat detailed level. The two cases vary to some extent in this activity since the situation in the computer game case is fictitious whereas the situation in the training and assessment system to some extent is given beforehand. In the case of designing multimedia systems for train-ing and assessment purposes, the author would try to obtain a deep and varied under-standing of the work situations that have to be demonstrated by the multimedia system. The multimedia author would primarily approach this by interviewing people in the work place or observing the same people in performing their actual work tasks.

... you learn a lot, and it is all about gathering information because you have to write about these figures, and make up figures where they can say – yes I know that guy, he is called Jens and he works in the warehouse ... Multimedia Author

However, he did also stress that the figures have to be anonymous in the sense that no individual in the organization should be directly identifiable as one of the figures in the story. He would also study communication and relations between people internally and between different departments. When trying to understand the culture and opinions of people he would normally look for any written material that is accessible like public relations leaflets, internal newsletter etc. As he expressed it:

... the page that contains complaints in the newsletter is often good, I will often sort the information there, but you can find some of the stuff people often think about ... Multimedia Author

In opposition to internal newsletters, public relation leaflets would often give a positive and partial picture of the company; however, it can be important in determining the signals the company wants to give its surroundings. In this activity, the game author would be faced with another situation since he is not trying to simulate a real world environment, but instead he has to define and describe a fictitious world where the narration can take place. In certain cases, the situation for the future game is partly given beforehand since the game will serve as the latest version of a series of games for children. The outcome of the activity is quite similar for the two kinds of interactive narratives. It is a description of situations where the narration can take place and description of characters involved in the situations. Also, the interaction and selections made by the future user is also described, as well as the manner in which these interaction and selections influence the narration told.

Last the story has to be written in greater detail and the idea is to write the manuscript in order to create lines and speeches for the characters in the narration. The multimedia author made a clear distinction between writing the story as a short story and writing the lines of the characters in the narration. However, this distinction often caused problems in the communication between the author/producer and the company.

... it is still difficult for the company to understand what is important in the short story ... sometimes they will actually correct words in the short story because they do not want the actual words to be said, however at this point it is not specified what the characters should say ... Multimedia Author

The game author stressed that the primary purpose of the manuscript is to make the lines of the characters more cogent, e.g. from whose viewpoint should this line be experienced by the user. Also, he has experienced that the manuscript has to be rather detailed in order to avoid any misunderstandings between him and the rest of the production team. Usually the manuscript will end his participation in the project and any ambiguity or misunderstandings may have severe consequences at this point in the design process.

2.2 Narrative Structures

As outlined above, during the writing of the short story and the manuscripts, the narrative or the story-telling part of the system is developed. The study revealed a number of commonly used narrative structures applied by the two authors. We have identified three, in our opinion significant, narrative structures with different characteristics and different levels of complexity. These structures played a significant role in the writing of the narration and they also played a major role in the control of the design process (also, they form the interaction of the systems).



Figure 2: Abstract narrative structure based on a single selection path

The single selection path structure, as illustrated in figure 2, is one of the simplest narrative structures for an interactive narrative (the most simple structure for an interactive narrative would be situations where the user has only one selection to choose from which we will not cover in this paper). At a given point in time, the user will be placed in a particular situation as indicated by A, B, C, D, E, or F. In each situation, the user will be able to make or required to make a selection in the application. Each circle in the figure corresponds to a situation, temporal or spatial, where the user is able to or required to make a decision. In situation A, the user has three selections to choose from where only one selection, S2, has been defined as correct. If the user nevertheless chooses to select either S1 or S3 the user will end up in the same situation forced to choose from the same pool of selections (the chosen one may in some cases be disabled). Between the selection of a wrong option and the return to the situation, the user will receive some sort of feedback on his or hers action, e.g. a full-motion video sequence, a sound, or a simple text message indicating that the user chose a wrong answer. According to the multimedia author, this narrative structure was often used in applications where the user should be trained in specific work tasks. As implied by the figure, this kind of structure will often be characterized as being highly temporal where situations are happening in a logical timely order and where you cannot normally go back to a previous situation (unless wrong). As a specific example, the multimedia author designed an interactive narrative for training (and assessment) of employees in a major shop. When an employee would use the system then at a given point in time, a customer will ask the employee a specific question, and the user is presented a number of different selections to choose from. In this situation, a wrong answer may be too impolite or too polite in the given situation. Every wrong (and right) selection is registered and used for later evaluation of the user. As a direct result, this structure is well suited for evaluation of practical specific skills and the multimedia author often applied this structure for training of skills of point-of-sale staff. On the other hand it is not well suited for assessment or training of more general skills, e.g. assessment of personal skills of managing an organizational unit of a company. Due to its overall low complexity, this narrative structure could be designed and described in a simple or general tool. The multimedia author would often use a word-processing tool for description of such a structure.



Figure 3: Abstract narrative structure based on multiple selection paths

The second structure, multiple selection paths, has a slightly higher complexity than the previous structure (see figure 3). The user is able to select different paths through the narrative. This implies that different users may have different experiences when they use the system. The basic assumption is that not all users will follow the same path thus experiencing the same situations illustrated by the system. In opposition to the single selection path structure, none of the presented selections are in reality more correct than others. As a specific example of this structure, the multimedia author had been working on a project for designing a multimedia assessment and training system for mid-level managers. The purpose was not to assess specific knowledge of the user, but more to assess general skills as operating manager in a company. Compared to the previous example for the single selection path on sales training, the user will in this case be faced with difficult and complex situations normally experienced by a mid-level manager. In a specific situation, the user, as the recently hired manager in a department, is having a staff meeting with employees in the department. The situation at the staff meeting is represented by full-motion video and at a point in time two employees start to argue. The system now provides the user with an opportunity to intervene by offering a number of selections to choose from. The selections are shown as small text-buttons on the screen. The various selections have different implications, e.g. the user can ask them to stop and request that the meeting continues or the user can ask the two persons about the problem. Based on the chosen selection the narration will continue on different paths, as in situation D1 where the user has three selections to choose from. The dotted line from this situation to situation F1 represents a timeout, meaning that if the user does not act in the situation within a given timeframe the situation will timeout and the narration continues regardless the lack of action.



Figure 4: Abstract narrative structure based on multiple exploration paths

The last structure, multiple exploration paths, is the most complex narrative structure of the three. Opposite the two previous structures, the multiple exploration path structure is not necessarily temporal in its character, but allows more spatial exploration in a narration. The game author for creating computer games for children often used this structure. A number of locations or environments is included in the narration as indicated by A, B, and C in figure 3. Within each location the user will often experience great freedom to interact or play. As shown within the location A, the user can move between a number of situations A1...A5. If the user is placed e.g. in situation A1, the user can choose from two selections leading to situations A2 or A3. The double arrows indicate that it is possible to return to a prior situation if necessary or wanted. Some of the arrows in the figure are shown as dotted lines, which indicate that they could be
conditionally active, e.g. you can only move directly from B3 to B1 if you have already found a certain piece of information. The game author gave an example of this narrative structure. He has been authoring a game for children called Cheese War in the Milky Way (direct translation from the original Danish title). The narration contains a number of characters such as the wizard called Bellini, a boy called Dennis, and a dog called Herman. The narration takes place in a number of different locations or environments like A, B, or C in figure 4. Within each location, the user can control and interact with the above characters and each location provides the user a great deal of freedom of interaction and navigation. However, only when certain conditions are fulfilled, the user is able to move from one location to another location. In the game moving from one location to another means moving from one planet to another planet. Also based on the previous selections and actions of the user, different things will happen at the new location.

... for instance you have an inter galactic cheesemonger who trades on the Earth as a starting point, later he can go to other planets and based on his previous record, different sorts of things will happen on his way ... Game Author

This narrative structure is applied in many computer games, e.g. action games and adventure games. Action games are often characterized by the fact that you can only move from one location to another location if certain conditions are exactly fulfilled, e.g. that you have found a specific key to open the door to the next level and thereby move to a new location. The game developer believed that the rigid fulfillment of conditions limits the possibilities for creating an interesting and catchy narrative. He would rather design and create a more open narration where different options are possible. This implies that you can play the game more times and experience something new and different each time.

3. COMBINING INTERACTION AND NARRATION

The previous section gave some illustrative examples of different kinds of narrative structures applied in interactive narratives. This section discusses these structures and tries to identify solutions to the problem of interactivity and narrativity.

3.1 Interaction and Narration

As mentioned above, the game developer desired to create a more open narration with more options, i.e. there should be a next step beyond the multiple exploration path as illustrated in figure 4. The question is: does this make sense, and, if it does, how can this be achieved?

In opposition to the descriptive genre that structures space, the narrative genre structures time. Normally, narratives build on a conflict between the protagonist's intentions and the resistance offered by his surroundings. According to Ryan [23], the basic tension is between the actual world and the possible worlds of intention, obligation, knowledge, and desires. For example, the actual world may not match the intentions of the protagonist who embarks on a quest to set things right. Additional tensions can be generated by conflicting possible worlds: conflicts between duty and desire (the classical French tragedy), or between obligations and knowledge (Oedipus). Bremond who defines the basic narrative unit as a modal progression from potentiality via

actualization or non-actualization to success or failure supports this analysis [4,5].

A skillful narrative will oscillate between these modalities: as Hitchcock knew, the possible is much more scaring than the actual:

...there is no terror in the bang, only in the anticipation of it.. Alfred Hicthcook [16]

The author of a narrative therefore needs to control time as well as modality. However, interaction means that the user too has acquired control of time, in the form of the selections he can make as illustrated by the three narrative structures. In this sense, will the game developer not defeat his own purposes when he desires a more open story with more options, since the more the user can control time, the less will the game developer be able to do it? Bordwell's clues and setups from the introduction will avail nothing, if the user chooses a path that does not contain the corresponding payoffs. Possibly some of the problems lie in the narrative creation process described in section 2.1, where the linear short story was found to be an important design document. Should the basis of the non-linear interactive narrative really be a linear story?

Let us apply a complete different perspective. The task of the author is not to provide a narrative structure at all, but merely this is the task of the user. As Crawford puts it:

...this artificial distinction then vests total control in the hands of the artists, and none in the hands of the plebians. I would ask, are the plebians so stupid, so dense, so utterly lacking in artistic sensibility, that we cannot afford them some measure of artistic control? The fact that some people are more artistically advanced than others does not argue for total control on their part, only control at a higher level of indirection... Crawford [10]

This "control at a higher level of indirection" could consists in constructing a world in which it is easy for "plebians" to generate interesting narratives. The artist is a world-maker, not a story-teller. In this case we would have solved the problem of the double control over time: the user has that control, and the author now controls the conditions for the user's control. This is also supported by two other experienced game authors, Rollings and Morris, who claim that games should provide the user with the opportunity to create their own stories [22].

This is clearly not the normal way of constructing narratives, but there are a number of precedents. One of the most successful ones is the work of J.R.R. Tolkien, cf. [26] He wrote two novels, *The Hobbitt* and *The Lord of the Rings*. But in addition, he provided a world of the kind we are looking for, *The Middle Earth*, which many other authors have used as a model for their books. The Middle Earth universe was built up long before the first novel was written and continued to develop throughout Tolkien's life. However, the notes were only published after his death under the title of *Silmarillon*:

Not only, however, does The Silmarillion relate the events of a far earlier time than those of The Lord of the Rings; it is also, in all the essentials of its conception, far the earlier work.[...] it became long ago a fixed tradition, and background to later writings. Tolkien [26] Maybe the future of interactive narratives will see authors construct Silmarillons that users can exploit to generate interesting narratives? If this is true, then an interesting question arises: can we characterize those worlds that are particularly good for generating narratives? However, before we begin to answer such questions, we should consider the issue of genres.

3.2 Interaction and Genres

Interactive narratives do not form one single genre. For example, training of skills requires the author to exert much more control than in systems designed to teach more general attitudes (Karin Levinsen, Courseware, personal communication). Skills can be imposed from the outside; attitudes need to come from within. For this reason, we cannot claim that more interaction yields better systems.

Another important qualification concerns the target of the interaction. Wibroe *et. al.* uses the distinction between story and plot from Section 1 to distinguish between two main types of interaction *story* and *plot interaction* [29]. The story is the fictive sequence of events as they unfold in the fictive timeline whereas the plot is the sequence in which information about the story is presented to the user. Usually these two timelines are different. For example, in detective stories, the murder is the first event to happen, but the last one to be fully told. The reason is that detective stories are based on a conflict between the possible world of knowledge and the actual world. Both detective and reader lack knowledge, and the novel is about the manner in which this lack of knowledge is liquidated.

Most games allow the reader to influence the plot, but the not story. In the action games described above, the user can move his character around in a landscape, collect information and gather the necessary tools. But it is information about a given story, and tools to solve given tasks. The basic story ticks inexorably away behind the scenes. Thus, although the interaction with the plot is high, the user may not "feel interactive" at all, since he cannot alter the underlying story. How can we make the story interactive? World making may provide the answer.

3.3 World-making

In the specification of worlds we have to change the notions in figures 2, 3, and 4 where the basic nodes are *situations* and the transitions between nodes represent user selections. In a world definition, the basic units should be *actors* and their *methods of interaction* — a well-known set-up in object-oriented analysis and design. In fact there is a genre of games that actually use this method of design, namely simulation games like SimCity.



Figure 5. Actant model of the Lord of the Rings

A possible candidate for a specification of a world (a "design pattern for narrative systems") is what is known as *thematic*

analysis in text analysis. A thematic analysis backgrounds the temporal structure and describes the themes and conflicts as they unfold everywhere in the work. A popular example of a thematic analysis is the actant model illustrated in figure 4 that represents one of the conflicts in *The Lord of the Rings*. It focuses on Frodo's quest of destroying the Ring and describes the relations of the main actors to this event: Frodo wants to throw the ring into the fire of Mount Doom, Sam helps him, and Sauron strongly opposes. Gollum is divided between opposing Frodo and claiming the ring for himself, and helping him to get rid of it.

Figure 5 is a thematic analysis since the split personality of Gollum is depicted as a single state containing contradictions. However, the two terms of the contradictions are temporally distributed over the whole novel: Gollum starts as the owner of the ring, looses it, and sets out to regain it. He is divided in his intentions until the end, sometimes helping Frodo, sometimes looking for ways of betraying him. As Sam says, Gollum is not to be trusted. However, when Frodo claims the ring himself, Gollum grabs it, but tumbles down into Mount Doom where he and ring is finally destroyed. Thus, Gollum ironically helps Frodo when he intends to oppose him.

But figure 5 does not tell us very much about how to design a computer system to generate interactive narratives! One obvious idea is to collect at set of interesting events and actors and implement the events as methods that take actors as parameters. But if Hitchcock is right in claiming that there is no terror in the bang, only in the anticipation of it, we get into trouble. What he is really saying is that it is not the execution of the method, but the process of passing parameters to it that is interesting! This is certainly true of The Lord of the Rings. The main theme of the book is: who is going to claim the Ring? In the beginning we learn that Sauron will, but as the story unfolds, the following persons are tempted: Gollum, Boromir, Saruman, Gandalf himself and in the end poor Frodo. Thus, we have a verb claim whose object is fixed throughout the story, namely the Ring, but whose possible subject changes, and it is the changing probabilities of actors to become the subject that drives the whole narration.

If we used a naïve approach, we would need to write something like this in our design: Claim (Sauron, TheRing), Claim (Gollum, TheRing), Claim (Boromir, TheRing)..., not meaning that the Claim method is executed, but merely that it has a chance of being executed with these varying parameters. This is clearly nonsense.

One way of making sense of this idea is to design the verbs as objects, and associate a subset of the well-known thematic roles to them [19]: Agent, Experiencer, Force, Theme (objekt), Result, Content, Instrument, Beneficiary, Source, and Goal, plus the circumstantial roles of Manner, Time and Location. If roles are objects too, each verb-object can be associated to one or more of the role-objects. Frodo's quest, throwing the Ring into Mount Doom would be represented as an instance of the quest-object to which an agent and destination object is associated. Parameter passing is represented by glue-objects that associate actors to thematic roles. The glue object represents the probability of the actor filling that role. Possibly three factors should be included: obligations, desires and capabilities. In the example below, the obligation part of the glue is strong, the capability part very weak:

I will take the Ring, he said, though I do not know the way Tolkien [25].

The glue-concept enables us to represent the fact that several actors strive to fill the same role. For example, the *claim*-verb takes an agent and a theme, and its theme is always The Ring, but many actors glue to the agent role (Sauron, Gollum, Boromir, etc.). The suspense of the book lies in the way the strength of this glue changes. From this follows that the effect of the events is to change glue! Sauron wants to *take* the ring from Frodo (role = source), and the effect of his attempt to *capture* Frodo is therefore to increase Frodo's ability to play the role of source for the *take*-verb. Similarly, Frodo's attempt to hide from Sauron aims a decreasing his chance of playing this role.

Whereas physical events influence the capability component of the glue, communicative events influence its obligation part. An example is the meeting in Rivendell where Elrond Halfelven *appoints* Frodo to be the ringbearer.

In section 3.1, we quoted Ryan [23] for the thesis that good stories are based on conflicts between the actual world and the possible worlds of desire, obligation, intention and knowledge, or internal conflicts between the latter. These conflicts can now be represented by the components of the glue-objects. For example, Frodo suffers from the conflict of desiring to take on the ring, and a prohibition against doing so:

He felt that he had from now on only two choices: to forbear the Ring, through it would torment him; or to claim it, and challenge the Power that sat in its dark hold beyond the valley of shadows. Already the Ring tempted him, gnawing at his will and reason. Wild fantasies arose in his mind.... Tolkien [25]

This can be represented as a strong positive desire component and a strong negative obligation component of the glue binding Frodo to the agent role of the verb *take on*.

Four other provisions are necessary to get going:

- 1. An event has a modal value that is calculated on the basis of the glue of its thematic roles.
- 2. An event executes when its modal value is above a certain threshold.
- 3. Changes of the modal value of an event must be staged.
- 4. The execution of an event must be stored together with information of its relations to other events at the time of execution.

Point 3 is crucial: we must be *told* of the changing odds of Sauron claiming the ring, and the system must display Frodo's agony when obligation strives with desire, even if Sauron never succeeds in claiming the ring, and Frodo never take on the ring. Changing variables are no good if they are hidden to the user.

Point 4 reminds us that the execution of an event may not be presented to the user at the time of execution, since we may sometimes need to have different timelines of story and plot. If the execution of an event is presented later than storytime, the system must contain a record of it that can later be retrieved and displayed. At this point the reader may have grave misgivings about the practical relevance of our analysis. However, the ideas behind Crawford's *Erasmatron* engine are not very far from the ones presented above:

...an Event is like a sentence; it is a record of one actor doing some verb to another actor...There's a lot of information stored in an Event, but the core information is short and sweet: subject, verb, and direct object. An Event will always have these three components...Crawford [11]

Figure 6 summarizes the above ideas (shown with the UMLnotation). There are four main classes in the system, *actants*, *events*, *glue* and *roles*. An actant is associated to zero or more glue-objects, whereas a glue-object is associated to exactly one actant and exactly one role. A role-object may associate to zero or more glue-objects (potential fillers of the role) and is the property of exactly one event. Events affect glue-objects, and contract "narrative relations" to other events to be used in the record (examples: an event triggers another event, or is executed to prevent another event). Finally, we need to represent the emotional relations between actants (the *sympathy* association in the diagram).



Figure 6. UML diagram of a possible narrative system.

If a system is specified in this way, the story would be interactive too for the very good reason that the system contains no explicit representation of a story at all! It only contains possible events and interactions associated to the actors of the story. The narrative itself will be an *emergent* phenomenon; it will emerge when the actors are set in motion in a given actual world (a phenomenon described in [22]. The purpose of the design process is to select a set of possible events and interactions that will often let interesting narrative emerges. Tolkien knew how to do this. Why not us?

4. CONCLUSION

Research has found the design of interactive narratives to be a difficult and challenging task. The introduction of narration as a key element of the future system and the presence of a high degree of interaction with the system makes it hard to design and predict the use of the system. The inherent problem of integrating interactivity into narrative structures is difficult.

Prototypical narrative structures were presented as illustrative examples of how to combine interactivity and narrativity. The narrative structures were identified in an empirical study involving interactive narrative authors and three narrative structures from an empirical study with different degrees of interactivity were introduced. A major problem in interactive narratives was identified, namely that both user and reader need to control time. Therefore a fourth step in complexity was suggested: in this method, which is very close to ordinary object-orientation, the designer creates a world model of a special kind which the user exploits to generate interesting narratives. This method will still have concept formulation but will supplement narrative structure design by world design and experimentation (cf. similar ideas in [22]).

ACKNOWLEDGEMENTS

We would like to thank the two authors of this study who volunteered in the interviews. We would also like to thank the anonymous reviewers for valuable comments on earlier drafts.

REFERENCES

- Benford, S., Bedreson, B. B., Åkersson, K.-P., Bayon, V., Druin, A., Hansson, P., Hourcade, J. P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K. T., Stanton, D., Sundblad, Y., and Gustav, T. (2000) Designing Storytelling Technologies to Encourage Collaboration Between Young Children. *Proceedings of the CHI 2000 conference on Human factors in computing systems*, pp. 556 – 563
- [2] Booch, G. (1994) Object-Oriented Analysis and Design with Applications. Benjamin/ Cummings, Redwood City, California.
- [3] Bordwell, D. (1985) *Narration in the Fiction Film*. London: Methuen.
- [4] Bremond, C. (1966). La logique des possible narratifs. *Communications* 8: 60-76.
- [5] Bremond, C. (1970). Morphology of the French Folktale. *Semiotica* 2: 247-276.
- [6] Brooks, K. M. (1999) Metalinear Cinematic Narrative: Theory, Process, and Tool. PhD Thesis, Massachusetts Institute of Technology, MediaLab
- [7] Christensen, A. S. & P. S. Jensen (1996). A multi vessel training system for highspeed craft. *MARSIM* '96, Copenhagen, Denmark.
- [8] Crawford, C. (2001) See http://www.erasmatazz.com/library/Lilan/inimical.html
- [9] Crawford, C. (2001) See http://www.erasmatazz.com/userdocs/engine.html
- [10] Crawford, C. (2001) See http://www.erasmatazz.com/library/Lilan/inimical.html
- [11] Crawford, C. (2001) See http://www.erasmatazz.com/userdocs/event.html
- [12] Eriksen, L. B. and Skov, M. B (1998) Evaluating Software Engineering Modeling Concepts for Interactive Narratives Design. *The 21th Information Systems Research Seminar In* Scandinavia: Information Systems Research in Collaboration with Industry, Sæby, Denmark, 1998, pp. 185 – 195
- [13] Eriksen, L.B., Skov, M., and Stage, J. (2000). An Empirical Study of Two Training and Assessment Multimedia Design Processes. *The 1st International Workshop on Narrative and Interactive Learning Environments*.

- [14] Gorriz, C. M. and Medina, C. (2000) Engaging Girls with Computers through Software Games. *Communications of the* ACM, 43(1), January, pp. 42 – 49
- [15] Grice, R. (2000) I'd Rather Play Computer Games Than Do Real Work. Wouldn't You? *Make It Easy 2000*, IBM
- [16] Hitchcock, A. See http://nextdch.mty.itesm.mx/~plopezg/Kaplan/quotes/quotes. html
- [17] Holbæk-Hanssen, E., Håndlykken, P. and Nygaard, K. (1975) System Description and the DELTA Language. Norwegian Computing Center.
- [18] Jacobson, I., Christerson, M., Jonsson, P., and Övergaard, G. (1992). Object-Oriented Software Engineering. Addison-Wesley, Wokingham.
- [19] Jurafsky, D. & J. H. Martin (2000). Speech and Language Processing. New Jersey: Prentice-Hall.
- [20] Kolstrup, S. (2001) Narratives: Different Narratives, Different Theories for Different Media? In Qvortrup (ed) Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds. Springer Publishers, London, pp. 275 – 291
- [21] Konzack, L. (1999) Edutainment for Girls: Analyzing Barbie Fashion Designer. In Nulden, U. and Hardless, C. (Eds) Nordic Workshop on Computer Supported Collaborative Learning. April 22 – 25, Göteborg, Sweden, pp. 47–56
- [22] Rollings, A. and Morris, D. (2000) *Game Architecture and Design*. The Coriolis Group.
- [23] Ryan, M-L. (1991). Possible Worlds, Artificial Intelligence and Narrative Theory. Bloomington & Indianapolis: Indiana University Press.
- [24] Skov, M. B. and Stage, J. (2001) Using Software Engineering Approaches to Model Dynamics in Interactive Software Systems. In Qvortrup (ed) Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds. Springer Publishers, London, pp. 404 – 421
- [25] Tolkien, J. R. R. (1974). *The Lord of the Rings*, I-III. London: Unwin Books.
- [26] Tolkien, J. R. R. (1977). *The Silmarillon*. Georg Allen & Unwin: London.
- [27] Umaschi Bers, M., Ackermann, E., Cassell, J., Donegan, B., Gonzales-Heydrich, J., DeMaso, D. R., Strohecker, C., Lualdi, S., Bromley, D., and Karlin, J. (1998). Interactive Storytelling Environments: Coping with Cardiac Illness at Boston's Children's Hospital. *Proceedings of the CHI 1998 conference on Human factors in computing systems*, pp. 603 – 610
- [28] Webb, B. R. (1996). The role of users in interactive systems design: when computers are theatre, do we want the audience to write the script. *Behavior and Information Technology*. Vol. 15, no. 2, pp. 76 – 83
- [29] Wibroe, M. K. & K. Nygaard, & P. Bøgh Andersen (2001). Games and stories. In Qvortrup (ed.): Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds. Springer Publishers, London, pp. 166-181

POLYGON DESTINIES:

The Production of Place in the Digital Role-Playing Game

Julian Holland Oliver Interactive Information Institute RMIT University, GPO Box 2476V Melbourne Victoria 3001 Australia 61.3.99252306

nearhere@selectparks.net

ABSTRACT

Historically, the 'corporeal intentionality' and sensitive zonings of the body subject are pre-requisites for the production of (a sense of) place. Digital Role-Playing Games however, in a simulation of this inter-animation, have been highly successful in acheiving this effect through a body avatar with representative sensitivities. First, this paper illustrates how the production of place is integral to navigation and progression through a Digital Role-Playing Game. Secondly, it reads the work of phenomenologists, Bergson, Casey, Merleau-Ponty, Heidegger and Serres into a discussion of how a dematerialised sense of place is produced and operated. Third, and overall, this paper considers the Digital Role-Playing Game as foundation for a critique of failings in digital simulation practice, positing it as working example of the 'worldly transformation' of mathematical 'space'.

Keywords

Game, Gamescape, Phenomenology, Place, Spatium, Interanimation, Quasi-Object, RPG, MMORPG, Local, Morphology, Avatar, Geometric, Simulation, Cosmology

1. INTRODUCTION

Unlike the simulation, which attempts an acutely representative relationship with the real world, the RPG [1], while borrowing from those technologies of representation, is not as preoccupied with facsimile.

Though advances in simulation technologies allow for a rich emulation of the real; gravity, inertia, flesh etc, the RPG chooses to bend the rules in the ends of gameplay, somehow managing to include all the ingredients necessary to create environments capable of richly immersive experiences.

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

RPG's have a massive popularity as complete worlds, and as such are actively inhabited. But how can such a deep sense of place (a function of such presumably material conditions) possibly be produced within patterns of interactive pixels that loosely emulate realworld conditions, and why is the production of this effect so important to gameplay? Short of attempting tedius speculation as to why so many millions of people would rather be somewhere else, we can look to the RPG for clues as to what is the rubric or code of this thing we know as a place, being that it is so effectively reconstructed in these games.

As much as cultural investigation might provide sufficient reason for recent popularity in this genre, technical developments in both hardware and game design are as much responsible; todays gameworlds are not only richer in visual detail, but offer the sense of a full and expansive universe. On first appearances new gamescapes like those of N64's Zelda, and PS2's Final Fantasies appear more preoccupied with producing this world effect than providing an arena purely for combat, dextrous operation and problem solving. Players have responded to this design perspective; RPGs are widely talked about as complete worlds, as actual places. This is an impression many traditions within the representative arts have sought after for centuries.

2. POLYGON DESTINIES

Reduced to it's graphic constituents, the RPG, like all 3D games are comprised of millions of polygons, little coloured shapes each with three or more sides all collaborating to provide a rich illusion of space. The polygon comes from a long legacy of geometric science whose first appearance was in a model of the universe found in Plato's Timaeus, First comes the point, then the line, third the triangle, four the simplest non spherical solid - a triangular pyramid.

In the geometric universe, forms are described as a composition of points falling within the three primary vectors of x,y,z. The subject is itself a point [2], whose visual subjectivity is represented as an equation of scientific perspective or 'vanishing point'. This field of algebraic-analytical relations allows for the purely mathematical construction of any plane within three dimensions, and has come to be synonymous with the largely universal idea of 'space' as it is scientifically understood. The digital visualization of this spatial model coupled with the visual configuration of the interacting subject can supposedly provide emulation of any object or environment - even environments we know very much as places. Simulations can be found for the Vatican, the Empire State Building, Niagara falls, the White Cliffs of Dover, our entire continents are in the process of being digitally reconstituted as we stand on them.

Heidegger however, in his Building Dwelling Thinking says that the understanding of space as comprised of points (spatium) produces not objects or places but multiple positions. When the dimensions of height, breadth and depth are abstracted from space as intervals, space itself and everything in it becomes a complex of these analytic-algebraic relations within the manifold of the three dimensions. He says,

"The space provided for in this mathematical manner may be called "space," the "one" space as such. But in this sense "the" space contains no spaces and no places...only allowing for the computing of these magnitudes."

A universe of coordinates he says is a universe of positions. What lies between these positions, the interiorities (places) and their simultaneous subjectivities, is of no significance in this model. This of course sounds disastrous for the legacy of a virtual reality, how can that sense of really being there possibly be produced when the very medium of this high representation is innately dissolving of the experience of place itself.

Henri Bergson however astutely points out that geometric space is not so much an arbitrary construction laid over the world as a filter or gauze so much as being inextricably written into both speculative and deductive reasoning through the very composition of the body. He says logical thought itself has a basis within the geometrics of the body in action; as the most efficient course of action for the body. Therefore it is geometric thinking which persists, in that speculative thought is the process of defining outcomes for any course of action. Funnily enough the RPG is full of both speculative and deductive reasoning, the solving of riddles, life threatening choices made through the strategic navigation of options. So what are we left with now; when the very substrate of the gamescape knitted with a placeless and denuded geometric logic, whose very vehicle of progression is the work of dry and logical thought? The RPG sounds hardly the site for a rich sense of place, capable of wonder and investment!

Yet the RPG is already the most popular game in Japan, so popular that when Final Fantasy VIII was released so many children took the day off school and so many adults the day off work that the Government released a public statement that videogames would not be released on weekdays. The RPG is one of the best selling genres of game in the world. More computer games were sold in America last year than books, and already one in five households in Britain owns a PlayStation - a huge proportion of these statistics include RPGs. And if as these statistics suggest, that more and more people will continue to invest in these ephemera as places, then we have a new kind of landscape to contend with, one that asks us in turn to innovate both popular and historical conceptions of landscape, place and social space. But the RPG is a paradoxical landscape, in that being dematerialized it appears incapable of occupation. This attitude persists up until an RPG is first played. For it is through the very action of gameplay that gamers build into what is otherwise a mere plane of interactive pixels.

We can see a capacity for spatial imagination at work in all games, which are themselves always a spatialisation of some sort, if only of action itself. While chess is an obvious example of this it is useful in understanding how the basic interplay of boundary and territory may become an engaging platform for action. The seminal abstract game Tetris, while being without the third dimension provided a gruelling arena for time to compete with the management of area in an unnegotiable rain of bricks. Gameplay holds this territorial logic together in the tensions of it's internal competitions. Competition within this innate territorial logic relies on a perfection of ability within the computer game, which always gets more difficult as the game goes on. But too much of either challenge and reward always results in a bad game. This elusive margin for the constant and persistently engaging improvement of performance is the quest of all game designers, and what is strange is that when this margin is found it is so often found for millions of people. - though some were better than others, everyone found Donkey Kong perfectly difficult.

The polygon has allowed for this spatialisation and perfection of action to be oriented in a very particular way, namely by the representation of spatial illusion which simultaneously produces the effect of an environment. With the 3D game the manifolds of territory are immediately complexified. The addition of this extra dimension, brings a capacity for movement that inevitably activates a curiosity for exploration.

3. FATEWORK

Fuller and Jenkins, in their Nintendo and the New Travel Writing configure these representative universes of the RPG as an interactive story, or, in line with it's explorative nature, as a spatial story of the like found in renaissance travel logs or epic novels. However on closer inspection, in other words playing these games, it becomes apparent that the story comprises mostly a context for events and the culture of gameplay. Though the story may be actively played into existence in a classic RPG like Final Fantasies, it better understood as an annex and addition of a more basic procedural logic within game design itself. My interviews of gamers (and personal experience) reveal that for the most part gamers play RPG's to firstly explore and secondly try and make it through to the end. The emphasis on story in modern RPG design with games like Final Fantasies or Baldur's Gate grounds the gameworld in a cosmology that gives the occurrence of events real meaning. More so the RPG need not include any sole agenda of epic travel. While still being explorative Sega's immensely popular Shenmue is an urban RPG set in the 1980's

This story, if you could really call it that is told to you in the inevitability of game play, the story itself isn't interactive, the game is. An interactive story implies that there are many different stories or outcomes to be had, this is not the case in RPG's today. In this way the RPG shares very little with the

hypertext novel it is often likened to. A major limitation in developments in this area is that even if there were to be an unlimited amount of labour and funding to make an RPG with an extensive branching plot, the whole thing needs to be squeezed onto a CD or cartridge. In this way it is ironically the economy of information storage that ensures we cannot build infinite universes - which is exactly why it is so important they feel infinite. This economy of disposable disk space alongside the expenses of time and labour inevitably influence a particular approach to game design, where progression in the game operates as amostly linear stream of events. This event stream however (while prevalent in game design generally) surfaces in the RPG in a way that works the gamer into the worldspace, into it's places.

Because there is really only room for a continuum that diverts occasionally, each scene is constructed as the means to progress to the next, and as such become constructed within a model where progression into another scene can not occur until certain tasks have been fulfilled - the golden key effect. Sometimes this is enforced with an object as iconically literal as the stone door of TombRaider, other times, as in PlaneScape or Baldurs Gate, passage is barred because the character has not fulfilled certain designated events at an earlier stage and so must go back. This device is important in tying down the character to any narrative obligations or design agendas at work within the game.

Precisely because we must find our way out of one situation and into the next, RPG game events have the feeling, and operate, as a pre-eminent structure or fateline, one that pervades a special importance to being somewhere[3]. There's always another door to find, a message with information we'll need, an inconspicuous amulet that proves to be absolutely necessary later...

It is at this point that the RPG develops an intensity not capable in the story.

Now within the dancing patterns of pixels we have a message, it is as though these places and events were left there for us especially. Potentialised with a greater significance than others, they feel like evidence of a universal logic or order within the gameworld.

The universal logic I'm speaking of here is not necessarily strategically set-up to produce the player within a solipsistic universe, so much as being foundational to game design itself; there wouldn't be a game without the task of earning the right of way. That there needs to be a series of tasks to perform before being able to access another level means that objects and events have the feeling, when playing, of being grouped around you. As a result these objects and situations seem to be reflections or apparitions of some kind of intentionality or will of the gameworld.

Because of this dynamic association with the gameworld through the centralised operations of the user, we can think of objects and events within the gameworld as relational objects, or to use Serres's term, 'quasi objects'. These potentialised objects and situations organise not only other objects and situations within the gameworld, but inversely organise the movements of the player as well. More than just a symbol, the quasi-object, because it is a relational object, is written into a morphology that binds the universe of the gameworld together, both in time and through the user.

Edward Casey notes that places attract and gather, objects, languages, people, places and things and that these are held in a particular configuration and as such make a relational sense between them. These relations support this universal logic that manifests in play as a total atmosphere, an indescribable mood that is particular to and pervades all things in the game

RPG's are full of objects, situations and chains of events that work beyond the mere associative attraction/likeness of metaphor, but work in themselves and together to produce learnable operating systems that often transcend specific games themselves. Many quasi-objects are recognisable even across game genres, like the convention of the health pack. In operation these objects etc need not have any representative association with objects etc in the real-world. They only need to work as a landscape, as a weapon, or a health pack to soon become those things; what begins as a metaphor soon becomes the thing it stands in for (where the metaphor is a rhetorical association that hasn't yet been operated). When you're down to 1, the health pack is very real.

4. PERFORMING 'REAL'

"Symbols, or general signs, have become associated with their meanings by usage. Such are most words and phrases and speeches and books and libraries."

Charles Sanders Pierce

This works even where there is only a bare likeness, as in PacMan. However the signs of this game only suggest a situation that comprises of a maze, ghosts and two sizes of yellow dots. It does not promise to be any more than the sum of these components - that is its charm. The system of signs in a RPG however emphatically promise a worldspace through the depiction of geographical, anthropomorphic and social correlatives that when operated, in other words become related, produce a worldspace. In this way RPG's do not need to strictly mimic or simulate real world systems to effect the sense of working as a complete world. In fact it is often the inherent other worldliness of RPG's that allow for them to operate in their own magical way, with no need for explanation. This is true of most science fiction - at a certain point we must give in to the fact that the alternate world works on a set of rules incompatible with our own. This same resolve allows for game symbols to become working objects, taking the rest of the worldspace along with them. Because objects and events in the RPG are operable, they do not need to be activated and supported by the imagination, and so they are quickly forgotten as the stuff of fiction.

Fuller and Jenkins, in their, 'Nintendo and The New World Travel Writing,' however insist that these universal logics found in RPG's are one and the same with the story. Furthermore they qualify the RPG, as providing means for the player to restage the founding myths of rennaisance period of America, ultimately as a means of attaining symbolic control, They say, "The R.P.G not only allows players to identify with the founding myths of the American Nation but to restage them, to bring them into the sphere of direct social experience... an R.P.G takes children and their own needs to master their social space and turns them into virtual colonists driven by a desire to master and control digital space."

Aside from this problematic association with the American Founding myths - a great majority of games are made and played by the Japanese, and secondly that prerequisite of the player of the RPG is a child (Sony estimates that upto 20 percent of it's players are over 35 years) - we can question the assumption that any RPG is played purely to attain mastery of a landscape, in the sense of becoming a god figure. Mastery may be temporarily achieved, but the RPG gamescapes are never held and managed purely in the sense of the domination of territory as in a game like Civilisation II or the turn based battle games. The RPG gamescape ultimately refreshes itself ready to treat any player the same way all over again. We always begin at some point of entrance (or level), and in order of progression, from the known into the foreign. We know that these places are inhabited, and that the game will consistently and willingly respond, as though itself a single sensate entity; we enter these worlds superstitious, as animists, building as we go a system of object, site and event and learning it's rhythms. Because these landscapes will ultimately return to their original condition, occupation is privileged with having no responsibility to long term consequence, to history, as a record of these successes. More importantly, experience of this forgetful or refreshing landscape is cumulatively advantaged through repeat performance - we get to practice getting good at being somewhere, something not allowed for within the condensed thresholds of the urban landscape. In this way the RPG can be understood as a kind of psychological holiday within the thick of everyday life.

It is an incredible feeling to know a place in a game so well that every exacted movement or operation is requitted with precisely the response you anticipated. No longer is the game a complex of trigger operation and interpretation, the interval between perception and action, game and player, is closed along with any consciousness that this is all happening within illusory space.

Another way of putting, 'getting good at being somewhere' is becoming local.

Edward S. Casey in, "How To Get From Space To Place" qualifies place as both occupation and the production of knowledge from occupation; in other words it needs only to become local to become an authentic place.

"...precisely as surrounded by depths and horizons, the perceiver finds herself in the midst of an entire teeming place-world rather than in a confusing kaleidoscope of free floating sensory data. The coherence of perception at the primary level is supplied by the depths and horizons of the very place we occupy as sentient subjects. We come to the world - we come into it and keep returning to it - as places already there.....There is no knowing or sensing a place except by being in that place, and to be in a place is to be in a position to perceive it...Such knowledge, genuinely local knowledge, is itself experiential in the manner of lived experience..." The power very particular to the format of the game is that knowledge is aquired through a performance of memory. Taking Casey's own cue, we get 'from space to place' by learning how each representative situation or area works through operations within them. In this way, the local knowledge that is produced in a RPG is different from other interactive systems (like other games) because earning it is simultaneously the feeling of confident occupation.

For Casey, 'space' is a hypothetical universal originating in Euclidean geometry, it can only exist without beings because to put beings into a space is to immediately produce knowledge, which in turn produces place. In this way we can address the problem of Casey's prerequisite for the experience of place - how does the user get inside the gameworld, in the sense of being in that place? To resolve this well look at the avatar for it's representative sensitivities and, as a kind of dynamic suture.

5. ID-WARE

ID software's Castle Wolfenstein 3D is arguably the first fully 3D game. Gamers were amazed to find themselves looking out through the eyes of the agent within a fully rendered world seamlessly scripted to movements. This viewpoint has since been synonomous with the first person shooter style of game, often coupling the visual field with crosshairs in a cyborgian blend of vision and weapon. The RPG has a different phenomenology of optics, one that at first appears to be contradictory to the project of writing the user into the game experience.

From the action dominant diluted RPG forms like TombRaider whose intermittent filmic cuts and interactive camera tracking set up the gamer as a movie director, to the turn based, character heavy traditional RPG's like Squaresoft's Final Fantasy series, the gamers viewpoint is a floating eye following the character as they are negotiated through the gamescape. In several lighter strains of RPG like Zelda and TombRaider, the user eye shares most of the perceptual horizons of the character but is higher up, often looking over the shoulder or sometimes out and around the avatar occupying an advantageous perspective of the situation. The more classic RPG's like Planescape, Baldur's Gate and Final Fantasy posit the user perspective in a priviledged top down view of the situation, much more is revealed of the surrounding environment. Both these view points seek to spare no details of the characters interaction with space, it is about an absolute exposure of the operations of the avatar in the environment, and as such all the threats to the well being of that character are transposed into a body oriented vulnerability. We watch these characters be maimed, tire and die and as such they becomes the subject of a covetous or protective gaze, simply because we can see them. The first person shooter viewpoint is perfect for shooters like Sierra's Half-life or ID Software's Quake for this reason; because we can't see our representative bodies our own our representative death is less significant, so much so that when playing some first person shooters ones numerous deaths become about as significant as a paper-cut. This viewpoint allows the pure and undistracted engagement in combat. Deus Ex though being a great game, claims to be an RPG while ambitiously operating within the first person optic. But because we never get to see ourselves in the game3, roleplay must be heavily

supported by the imagination to compete with this first person interface for action, and investment in ones chosen role is more difficult.

The RPG asks that the player steps from these more utilitarian relationships with the avatar and into a relationship with a character - a kind of avatar that embodies not only a greater portion of representative humanness, but in doing so widens capacity for empathic investment.

In an RPG the character plays an essential role in the milieu of the game and as such is less preoccupied with combat than with an excellent adventure. Because of this, gameplay is of a slower pace and the life expectancy of the character is generally greater allowing more time and more vehicle for investing in their representative lives This intensifies as we move into the more traditionally derived (pencil and paper based) RPG strains like Baldur's Gate and Final Fantasy. Instead of looking through the body into the gamescape, the avatar of an RPG is configured as the subject of a reflexive and responsible gaze that is forever assessing possible courses of action for the avatar as we coordinate it in relationship to the events of any given situation. It is the gamers projected empathy and felt responsibility for the representative sensitivities and mortality of the character that produces an extended Cartesian reflex. This active relation constitutes a kind of cybernetic selfhood that in turn binds the gamer with the character through a foldback of both real and representative sensitivities.

Now, with this avatar/user bind in place, we can see how an RPG satisfies Casey's prerequisite of being in that place.

The avatar of an RPG is not a visitor of the gameworld, so much as a denizen of it both in milieu, and innately in it's internal construction. Numerous basic social exchanges produce and affirm the subjectivity of the avatar, weaving it into the milieu of the gamescape as a life existing amongst other lives. This subjectivity exponentially increased with the advent of network gaming and multiplayer, where other real time players interact with the user. But the avatar is written into the game on a more innate level.

6. INTER-ANIMATIONS

The very logic of game design ensures that the representative body of the avatar is already capable of completing any challenges that the game has in store. The avatar is ultimately the most capable of competing with the challenges ahead of them, a part of the world through what they can do in it. It feels this way from the moment we begin playing because we know that the gameworld was made entirely for us. As an agent a priori of the universal design, they are even more than just a denizen. These factors ensure that the players decisions are transported into the game, beyond the pure intentionality of the first person shooter game. They operate the mechanics, myths and logics of the gamescape through a fleshed out subject of that world, and as such the world feeds back through that embodiment, in turn animating the gamer. It is this feeding back of the world through the gamer that is contiguous with Kant's seminal theory of interanimation where place and the subject are autopoetically engaged in a continuous and dynamic reciprocation - a necessary factor for Kant in the production of place.

If as Casey says, that place has openings, thresholds, sensitive zonings that mimic the flesh of the body, then the same is true for the game. There is a shared pneumatic fleshlike structure between avatar and gameworld; the digital flesh of the body/avatar shares the same flesh of the gameworld. But this is not enough to produce a sense of place. The body must move within the place, as a dynamic component of it's flows. And Casey agrees, by noting that Galilieo's configuration of the body, as a punctiform object committed to the laws of gravitation and physics and without it's own self motion, is not capable of producing place.

Casey asserts that, "precisely by allowing us to make diverse entry into a given place...the body insinuates itself subtly and multiply into encompassing regions". In his essay, 'How to get from Space to Place', he looks at Kant's bilateral body. With brachiations and multiply articulated structure, it has multiple and extensive simultaneous engagements with space, producing a bodily subjectivity as an extension of space > place. This also works as a means of place interfacing back onto the body through multiple sensitive points. Of course the representative sensitivities of the avatar cannot compete with that of the gamers own, even though many gamers will tell you they feel every pain of the avatar. But on a more intrinsic level however, in the sense of the bodily structure, the avatars subjectivity very much a part of it's place, written into the gamescape as an extension of it's symmetries and distances.

But what is special about the digital body of the RPG avatar, is that it's very structure, by being designed as ultimately perfectly fit to operate in the game world it already contains the necessary intervals of the gameworld, a perfect subject of it's physical dimensions; of it's anatomical challenges. For instance, of all the jumps we cannot make in a game, <u>there is always a jump we are meant to make</u>. In other words where ever the avatar can operate, their body is best suited to operate - and so often it is what we cannot do that leads us to what we are supposed to do, <u>a perfect and reliable reciprocation of effort</u>. Each ladder, river, dungeon, car park is made for the operations of this body in particular, even if against it. The avatars body is direct expression of their environment, written into the gamescape as a capacity for it's distances.

It seems we have come back to Bergson's body of intervals, from which we project possible courses of action in the gamescape. Looking at Lara in action we can see just how much an avatar can contain the geometry of the world in it's very intervals. Each movement fulfills a dimension in the gamescape.

These environments, as a geometric precipitation of the intervals of the avatar's own body works as the pure transmission of the geometric intellect, as a total manifestation of the Euclidean ideal we are supposedly so innately best suited to. Maybe this why operations in a gameworld can sometimes feel so fundamentally reassuring...

But though the avatar is so very much part of the gameworld in these ways, a body and a place tend to present themselves as particular; eg my body in this place. The body of the avatar is claimed by the gamer as an extension of this assertion. and so by way of this relation the gamer is equally frustrated by the avatar's containment. This dynamic exchange of real and representative sensitivies in the frustrations and competitions of gameplay facilitates an exchange that oscillates in and out of awareness of the iconographic individuations of avatars like Link or Lara - we sway between roleplay and becoming.

It takes a lot of work to get through a game - finding a way to the next place is always a significant event...And in an RPG often it's not what you do, it's when or how you do it. There's a scene in TombRaider III where the gamer has to find a switch that turns the whole room upside down, to allow access to a tiny window that's impossible to reach the other way up. It takes most gamers days to work that one out.

If there's one performative relation denominative to most videogames, it is the determination and overcoming of boundaries through gameplay. These boundaries describe the limits of experience in a game while also doubling as a kind of perceptual horizon. To open up another view or access another area is to open up a new field for action. In other words it's through the transgression of these boundaries that consecutive places reveal themselves to the gamer, the world exists before we get there, but we've got to get to know it to get through it.

As we saw earlier, place in an RPG is performed, and through the boundary, further places are performed into existence. It is this reflexive exchange, operating through memory, of the difference between environments, that produces the worldspace as potentially infinite. It does not need to be infinite, so much as the effect of it contained within the logic of the very action of progression. And, because we so often group, name and remember places by what we do in them, to play RPG's is to accumulate numerous secret histories, atlases and a sense of having lived in another world. These are histories and experiences that other gamers understand, where they can say, "you know that place where..."

7. FROM SPACE TO PLACE

The gamescapes of the RPG come as a result of early developments in the stepsister industries of virtual reality and simulation, yet don't seek to emulate the real. While the RPG might borrow from the technical advances of these practices, the agendas of the polygon are decidedly redirected into the production of other worlds. Polygons in an RPG collaborate as a skin for the production of a gamescape dense with signs, superstitions cosmologies and milieu. These are alternative ways of being in a world, alternative to the very logic that is innate in the substrate of their very make-up, that of polygons - the geometric unit of an ancient model that sought to measure and emulate the universe as accurately as possible. Ironically it is this conversion of the polygon that has redirected the founding

agendas of the 'digital space', reworking it in such a way as to be capable of personal investment, to become habitable - in short, a place.

8. Footnotes

[1] for purposes of this paper the terms RPG refers to a single player game employing an avatar undergoing a character development throughout the course of the game.

[2] the 'point' of the subject is often referred to in game play, developer tools and 3d modelling software as a 'camera'

[3] whereas in many other genres of game this sense of predestination is often carefully disguised in the interests of realism or other milieu / design agenda's

9. ACKNOWLEDGMENTS

My thanks to Chad Chatterton [for happy hours of conversation and gameplay], and Rachael Tempest, for both research assistance and commentary.

10. REFERENCES

Henri Bergson, Matter and Memory (New York: Zone Books, 1998)

Fuller and Jenkins, Nintendo and The New Wold Travel Writing (nettime.org, 1996)

E.Casey, 'How to Get from Space to Place...' Senses of Place, (School of American Research Press: New Mexico, 1997)

<u>M. Merleau-Ponty, The Visible and The Invisible</u> (Northwestern University Press: Evanston, 1968)

Martin Heidegger, 'Building Dwelling Thinking', Basic Writngs

The Edge Magazine, columns, Issues 82, 83, 89 (UK, Future Publishing)

Michel Serres, 'Gnomon: The Beginnings of Geometry in Greece,' A History of Scientific Thought, (Oxford: Blackwell, 1995)

Robert Markely, Virtual Realities and their Discontents, (Baltimore: John Hopkins University Press 1996a)

<u>Phillip Goodchild, Time and Evaluation</u>,' Antithesis - Special Ed, Time and Memory' (Melbourne, Depts. English and Cultural Studies, Melb University 1997)

Beyond Manzanar: Constructing Meaning in Interactive Virtual Reality

Tamiko Thiel Artist http://mission.base.com/manzanar/ Baaderstrasse 64 80469 Munich, Germany

tamiko@alum.mit.edu

ABSTRACT

Beyond Manzanar is an interactive 3D virtual reality artwork. It is written in VRML and runs off the hard drive of a fast gaming PC. It is shown as a room installation with the image projected life-sized on a large screen. One user at a time can navigate through the 3D environment in first-person viewpoint using a simple joystick; other can watch and share the experience.

The piece explores parallels between experiences of Japanese Americans imprisoned at the Manzanar Internment Camp in Eastern California during World War II, and Iranian Americans threatened with a similar fate during the 1979-'80 hostage crisis. The physical environment of Manzanar, strikingly similar to the landscapes of Iran, creates a poetic bridge between two groups linked by the experience of being "the face of the enemy."

Beyond Manzanar is an exploration of the capabilities of virtual reality to provide a strong dramatic structure and still involve users in creating their own narrative, adding a sense of participation and responsibility for the events that happen to them in the course of their explorations.

Keywords

Virtual reality, social uses of technology, dramatic structure, user experience, interactivity.

1. INTRODUCTION

The interactive virtual reality work *Beyond Manzanar* uses navigable 3-D graphics running on a fast gaming PC to simulate the Manzanar Internment Camp in California, USA, transformed by dreams and memories. The piece explores parallels between experiences of Japanese Americans interned at the camp during World War II and Iranian Americans threatened with a similar fate during the 1979-'80 hostage crisis. [1] A realist reconstruction of the internment camp becomes a framework for interior visions of personal responses to the betrayal of the American Dream, and imagined landscapes of Japanese and Iranian gardens explore the healing processes of memory and cultural grounding.

Beyond Manzanar was a joint project between myself, an American visual artist of mixed Japanese/German ancestry, and Zara Houshmand, an Iranian American theater director and poet. Both Zara and I had worked as creative directors at Worlds, Inc. in San Francisco, one of the first companies to bring out a commercial platform for real-time PC-based virtual reality. We were well acquainted with the technical capabilities of the medium from our commercial work and wished to explore the dramatic possibilities of the medium in a pure art piece.

We used theories of dramatic structure taken from theater, music composition and urban design to create an experience with a strong narrative structure. Within this structure, the interactive technology gives users freedom to find their own way through the environment and make their own discoveries. Their movements through 3D space translate them across time and cultures as the environment senses their presence and transforms around them. The first-person viewpoint and interactivity of the piece require users to take emotional responsibility for their decisions and the consequences. In the end, however, this is art and not life. We let the software loop to allow users to return and relive the spaces, now transformed into part of their own memories.

2. CONTENT AND BACKGROUND

Manzanar, an oasis in the high desert of Eastern California, was the first of over 10 internment camps erected during World War II to incarcerate Japanese American families solely on the basis of their ancestry. Though this specific instance was ruled "not justified" in 1988, mass internment of an entire group without due process "in cases of military necessity" is still legal. Ethnic groups whose countries of origin are considered rogue states by the American government can legally be interned without trial if tensions between the countries escalate into violence.

In 1979 during the Iranian hostage crisis there were physical attacks on Iranian Americans and calls to intern and deport them to Iran, regardless of their personal stances towards the hostage crisis. For an Iranian American, Manzanar is an especially ironic symbol of the threat of internment: the site itself is hauntingly reminiscent of the landscapes of Iran. The grid of roads drawn in the desert by the military echoes the geometric order of an Iranian paradise garden - a further irony, for the Japanese Americans did indeed create gardens, an ancient form of virtual reality, within the barbed wire fences. Manzanar is therefore the ideal symbolic site for a metaphorical exploration of parallels between the wartime experiences of Japanese-Americans and current discrimination against Iranian Americans as "the face of the enemy."

3. INSTALLTION AND INTERFACE

Beyond Manzanar is shown as a room installation with the image projected life-sized on a wall. This creates the feeling of an immersive virtual space while still allowing groups of people to view the piece together. It is written in VRML and runs using the blaxxun Contact 4.4 browser plug-in to Internet Explorer. Due to the large number of relatively large texture maps (256x256 pixels and some 512x512 pixels) it requires an AGP graphics card with minimum 32 MEG of graphic RAM and a fast processor with at least 256 MEG of system RAM.

One user at a time can control a simple, small joystick that moves the user in first-person viewpoint through the virtual space. There is only one path through the piece and the piece will loop indefinitely as long as the user keeps moving forward through the scenes. A user can leave at any time; the next user takes over the joystick and starts where the previous user left off.



Figure 1. Installation View

It was important to me to keep the interface and interaction simple, since an important part of our audience is former wartime internees. Even the youngest of them are now over 55 and the great majority have no experience playing 3D games such as Doom or Quake. I chose a joystick as the interface with the reasoning that it is the interface device used by handicapped people to steer motorized wheelchairs through real 3D spaces, and should be relatively easy for a variety of people to use in virtual space as well. I am delighted to be able to report seeing many little old ladies jumping at the chance to gain control of the joystick, and of course their youngest grandchildren and greatgrandchildren are equally adept, if not more so.

4. CREATING EMOTIONAL AFFECT

My personal technical interest in creating this piece was to experiment with the volitional and kinesthetic possibilities of VR, which for me meant that users should be free to explore and find their own way through the piece rather than be presented automatically with linear content or by such obvious affordances as a "next" button. I wished also to utilize the possibilities for full immersion and full identification inherent in a first-person viewpoint.

This creates a challenge: how could we give the piece a strong dramatic character while still allowing the users to construct their own narrative? The known weakness of using a first-person viewpoint, mentioned for instance by Andy Clarke and Grethe Mitchell, [2] is the difficulty of creating narrative interest when there is only a single camera viewpoint.

Neither film, theater nor dance theory provided me with insight into dealing with these problems. Theory from these fields tends to be character centered and deal with the development of dramatic tension between characters seen mostly in third-person viewpoint. Viewers see the work from a single fixed seat in a theater and the director therefore has quite a lot of control over what is visible in any given scene.

I surmised that with a first-person viewpoint, character development has to happen essentially within the users themselves as they explore the virtual environment. This means composing the piece from the perspective of the emotional states roused in users by their interactions with the piece itself. Brenda Laurel's work [3] provides interesting insights in both drama theory and its application to a more "human centered" form of interaction between human and computer, but I felt I needed a more abstract or fundamental theory that provided insight into composing the discrete units of the piece. I looked to the more abstract realms of time-based media theory: music and urban planning.

If music can provoke powerful emotional reactions and the feeling of deep meaning in the listener merely by arranging abstract notes in certain structures, it must be possible to do the same in any other medium as well. Leonard B. Meyer [4] addresses this phenomenon directly, proposing the thesis (to vastly simplify his argument) that music provokes emotional responses in listeners by arousing expectations and then playing with these expectations in various ways: fulfilling them, disappointing them, surprising them or by delaying their resolution. The composer creates a sense of *meaning* in the listener by bringing these various emotions to a resolution and conclusion by the end of the piece.

Focusing on the emotional states of a first-person protagonist was a familiar concept for me: my father, Philip Thiel, preaches this method in his work on urban environments, providing theory and tools for both analyzing and designing sequential urban spaces from the viewpoint of a person walking through the environment. [5] His analysis dissects an environment into the following components:

Space: the physical landscape itself (plains, mountains, rivers, etc.)

Place: the built environment created by human beings (buildings, roads, etc.)

Occasion: what people are doing in this place at a specific time (sitting in a cafe, holding a demonstration, etc.)

In *Beyond Manzanar*, the *space* is the Manzanar landscape, represented by its characteristic background mountain panorama. The space always remains the same but the *place* changes: sometimes an internment camp, sometimes a paradise garden. Each built environment affects not only what can be "done' in the space – the *occasion* – but also the user's emotional relationship to the space. The same mountain panorama produces a different emotional affect depending on whether it is seen framed by a paradise garden or by barracks and guard towers. This fact forms the basis of the dramatic structure of our piece.

In this paper I will refer to a specific combination of place and occasion as a *scene*. Each scene was designed from the perspective of the emotions that it should evoke in the user. Each scene has an emotional "ground note:' it should evoke either negative or positive emotions, for instance a feeling of alienation or of security. The sequence in which the scenes are concatenated is carefully arranged in the same way that a music composer arranges phrases to create dramatic structure within a piece of music, complementing or alternating emotional states, building suspense or releasing it dramatically. As the user becomes more experienced with the piece, looping through several times as many users do, the element of surprise is reduced but also replaced with the element of anticipation.

5. CREATING NARRATIVE IN A SCENE

Creating an environment or building in VR, whether an historical reproduction or a new fantasy, can produce an impressive stage set. For an artwork this is not enough, however. It must somehow be filled with the culture of those who built it, real or imagined; otherwise it remains an empty shell. If it replicates a real environment it should also use the characteristics of the technology to provide an experience that is impossible to get from visiting the real site. This for me is the excitement of virtual reality: it provides an excellent platform for creating gesamtkunstwerke in which an environment is brought to life by the inclusion of other material such as sound, images and texts (and for that matter other characters in the form of avatars.)

Thus the WWII camp scene in the piece makes extensive use of images and sounds from the Manzanar Internment Camp. The barracks are reconstructed from archival photographs and military data, but to emphasize their emotional effect and physical presence they are slightly larger than life-size, as adult internees tended to be very much smaller than the average height today. To emphasize that real people were involved and to give a sense of the conditions under which they lived in mass confinement, the windows of the barracks are filled with historic photographs showing daily life of the internees at Manzanar. To underscore a sense of personal loneliness, the only sounds are a woman's voice singing from the barracks and the crunch of ghostly footsteps – and of course the incessant, mournful desert wind.



Figure 2. Camp Life

To make it clear that this piece is not a documentary but a mixture of "what has been" and "what could be" we consciously emphasized the surreal to create the feeling of being in a dream. We found that virtual reality lends itself to visualization of figures of speech: in the sky above the internment camp, headlines from the war and anti-Japanese signs slowly fade in and out to literally fill the air with hate. It is important that the signs are not always there; sometimes the air is clear, which makes it that much more oppressive when the sky again fills with hate and foreboding.



Figure 3. Signs of Antipathy

Another way to increase the dramatic impact of a scene is to play cat-an-mouse with users and their emotional relationship to the space. The "quiet" scene where the piece starts (or re-starts) shows the empty landscape of present day Manzanar: mountain panoramas and a grid of roads through the desert scrub. The space seems unbounded and open, but if users try to "run away" a barbed-wire fence materializes in front of them. The physical object is enhanced with cultural content: the barbs of the fence are poems about exile written in English, Farsi (Iranian) and Japanese, some actually written by Japanese Americans in other camps, some from classical Japanese and Iranian literature. The user stands at the fence, confronted with a narrative voice that underscores the visual and physical message of the fence: You are a prisoner.



Figure 4. Rumi Fence-Poem

A feature of virtual reality that has strong dramatic possibilities is to play with the vertical dimension. One example in our piece is within the American Dream sequence, where the seemingly solid walls and floors of a room become transparent to leave the user hovering in the air above an internment camp. With this and all such technical effects it is important that they be used only when they reinforce the narrative of the piece, never merely for their spectacular visual effect.



Figure 5. Above the Void

6. DRAMA IN SEQUENCES OF SCENES

Scenes are linked by open doors or paths that function as portals into the next scene, an intuitive clue even for inexperienced users. In order to make it work, however, I had to make a harsh decision: once you go through, the door slams shut behind you. The first time this happens it is of course a surprise and a frustration. It is however also a very important part of the message of the piece: in real life, there is no going back either. You have made a decision and must live with the consequences, even if you did not realize what would happen.

Alternation between emotionally positive and emotionally negative scenes is the crux of the dramatic composition of the piece as a whole. In one scene, for example, users are inside a barrack but can look out the window to see the mountains of Manzanar framed by other barracks and the guard towers.



Figure 6. View Out of the Barrack Window

If the users then enter a Japanese-style room within the barrack, the barrack – and indeed the entire camp – disappears to be replaced by a pavilion in a Japanese garden. In the garden I used the traditional captured scenery technique to visually integrate the mountain panorama: the eye sweeps out over a pond, continues over the trees and to the far mountains. In the context of the camp the mountains had formed a formidable barrier beyond the fence; in the context of the garden they lead the eye and spirit up and beyond, uniting the man-made garden with the timeless beauty of the natural world.



Figure 7. Japanese Paradise Garden

This sense of unity with nature is deceptive and temporary, a momentary dream in the dull reality of camp life. Once users begin to explore the garden, they will inevitably trip a trigger that throws them out of the garden and back into camp. The mountains return to their role as a formidable barrier on the horizon. The scenery stays the same, but the user's emotional relationship to it changes.

Users are left with the feeling of responsibility, however inadvertently, for having destroyed the dream of the garden. There is no going back, the memory of the garden must suffice as the users move forward to confront the next scene. They are forced to be complicit in their own destiny – a favorite trick of sadists everywhere.



Figure 8. Fall Back into Camp

7. SUMMARY AND CONCLUSION

Traditional theater can integrate objects, images, sounds, texts and characters into a gesamtkunstwerk. Virtual reality can go beyond theater not only by exploiting the irreality of a space with no physical laws, but also by use of a first-person viewpoint to bring the user's own body and personal character into the piece. Projecting the image life-sized on a screen brings the user a physical sense of space and scale; allowing the user to navigate through the space on their own adds the kinesthetic sense of movement, orientation and place. Interactivity, whether through positional triggers or active manipulations of objects, makes users complicit in the dramatic structure of the piece, adding a sense of participation and responsibility for the events that happen to them in the course of their explorations.

The artist must understand the emotional effect of each of the virtual spaces on the user, and must organize these scenes into phrases or movements, sequences of scenes that can be combined to create a sense of the classic build up, climax and denouement used in the dramatic structure of theater. If the piece loops, every scene must have enough material to bear several investigations. This insures that each time through the piece users will construct a slightly different narrative, viewing and interacting with the material in a slightly different way, discovering new objects or seeing the same things in a different light as their previous experiences have now become part of their own memories.

According to Leonard Meyer, it is not enough to provoke emotional states; one must also provide meaningful and relevant resolutions in order to produce an emotional release of the tensions built up within the piece. Although the looping structure of *Beyond Manzanar* provides no true beginning or ending, it still has a strong dramatic climax followed by resolution.

In the "war" scene, Zara played on fears of Iranian Americans that the American bombers used in the Gulf War against Iraq could also be used in some future conflict against Iran. Users are swept up out of an Iranian paradise garden by an F-15. Tumbling in the air, they see the garden now framed by an internment camp – and by the sights of the F-15.



Figure 9. War

The F-15 finally passes, leaving users spiraling through the air high above the landscape of Manzanar. The "Resolution Poem," written by Zara specially for this piece, then appears against this backdrop. It asks the mountains, winds, earth and sky to bear witness to the history of this place called Manzanar, so its story may never be repeat.

Although this piece refers to events and people in the United States of America, we feel that our message is a universal one, applicable to any country anywhere in the world where politicians single out groups as convenient scapegoats in times of crisis. We hope that the experience of *Beyond Manzanar* goes beyond visual and intellectual stimulation to affect users emotionally, communicating to them a bit of what it means to face discrimination solely on the basis of belonging to a group with "the face of the enemy."



Figure 10. Resolution Poem, 1st Verse

8. ACKNOWLEDGMENTS

Many thanks to the sponsors who made *Beyond Manzanar* possible: Production was funded by an artist-in-residency grant from the Gifu International Academy of the Media Arts and Sciences, Japan. The project was developed on a Pentium® III processor-based PC, courtesy of Intel Corporation. Generous technical support is being provided by blaxxun Interactive, Inc. Initial research was funded by a grant from WIRED Magazine and the Asian American Arts Foundation of San Francisco.

9. REFERENCES

[1] For a well-linked site on the WWII internment of Japanese Americans see the San Francisco State University webpages http://bss.sfsu.edu/internment/documents.html. For parallels between the Internment and treatment of Iranian nationals during the Hostage Crisis see for example P-I News Services, "Carter Cracks Down: Iranian Students Face Deportation" and "Carter Act Stuns Iranian Colony," Seattle Post Intelligencer, Nov. 11, 1979, pp.A-1, A14, and Grabowicz, Paul. "Acts Against Iranian Students Disturb Japanese-Americans," Washington Post, Nov. 23, 1979, p. A21. In an incident that could only happen in real life, Japanese American Senator S.I. Hayakawa (who was a Canadian citizen during WWII and was not interned) took it upon himself to introduce a bill to "imprison all Iranians living in the country as long as they were not U.S citizens." Fogarty, John. "Hayakawa to Seek Internment of Iranians,"

San Francisco Chronicle, March 12, 1980, p.1 and back page. For reactions to this proposal from the Japanese American community see Eastham, Tom. "Hayakawa's 'seize Iranians' plan attacked," *San Francisco Examiner*, March 12, 1980, p.?.

- [2] Clarke, Andy and Mitchell, Grethe. Screen Play: Film and the Future of Interactive Entertainment, http://www.kinonet.com/theory/dcc/dcc_paper.html, 2000
- [3] Laurel, Brenda. *Computers as Theatre*, Addison-Wesley Publishing Company, Reading, MA, 1993.
- [4] Meyer, Leonard B. *Emotion and Meaning in Music*, University of Chicago Press, Chicago, IL, 1956.
- [5] Thiel, Philip. People, Paths and Purposes: Notes for a Participatory Envirotecture, University of Washington Press, Seattle, WA, 1997.

Semiotic and nonsemiotic MUD performance

Ragnhild Tronstad University of Oslo Dept. of media and communication Box 1093 Blindern, N-0317 Oslo +47 22850412 ragnhild.tronstad@media.uio.no

ABSTRACT

Using performance and theatricality as metaphors to describe MUD adventure games this paper starts out by defining a difference between 'performative' and 'theatrical' acts within MUDs. Performative acts differ from theatrical acts in that they are effecting 'real' changes to the MUD environment, thus directly influencing the game in one way or another. It is argued that to understand the performative aspects of MUDs it is necessary to separate 'acts' from 'meaning', and to pay special attention to the former in order to avoid confusing the questing process (which should be our main focus of attention) with the experience of having solved a quest.

Keywords

MUDs, acts, quests, performance, theatricality, act/meaning, performatives/constatives

1. INTRODUCTION

In this paper I will describe two ways of viewing MUD performance: as either theatrical or performative. While the former is playing with meaning and signification, the latter can be said first and foremost to serve a nonsemiotic function in the game. 1

I will start out by explaining how I use and understand theatricality, and its relation to the 'real world'. Then I'll turn to performance. Very briefly I'll present some of the most popular performance understandings, before I define the one I will be using in relation to theatricality. After defining my theoretical position I will show how I use the two concepts performance and theatricality to point to a difference in the ways characters perform in MUDs.

My main object of research is an adventure oriented MUD situated at the Technical University of Berlin, TubMud. If nothing else is indicated, my examples are taken from this MUD. The principles presented in this paper will not necessarily apply to all kinds of MUDs, as there are many types, different in both structure, content and purpose. I think, however, that they could be applied to most MUDs of the adventuregamelike type.

2. THEATRICALITY & PERFORMANCE

2.1 Theatricality.

Theatricality is the process that makes it possible to transform an everyday situation into some sort of fiction: either in the theatre, done by actors; or on the bus, in the street, or at home - by me or you, as either actor or spectator. Theatricality can be to show things in a different setting than the one in which they are normally perceived. It can also be just to see things differently. It can be done intentionally, or unintentionally. A misreading can be theatrical, as it creates a breach in the perception of 'meaning'

As MUDs are textual representations of spaces, places, objects and beings, nothing is really seen in a MUD. All representations are - more or less vividly – imagined, and they are imagined differently by each player, according to his or her previous experiences inside and outside the MUD. The experiences called upon while interpreting inside the MUD we could call 'context' or 'frame'. Thus, spaces, places, objects and beings are framed differently by each and every player.

This is why MUDs should be seen as a fictional world I imagine by interacting with room-, object- and character descriptions. I know that there are real persons behind the characters, but I also know that unless I have met these people in real life, my interaction with them is more theatrical than real. I imagine them.

In this way, MUDs come close to being a theatre-for-oneself, as defined by Nicolas Evreinoff almost one hundred years ago. (Evreinoff 1927: 187-198) The theatre-for-oneself is a theatre not-for-everyone, and still everyone can experience it. It simply means to transform what we know as 'reality' into a theatre, by pretending it is staged. This change of attitude towards something familiar produces a change in perception: Signs are reloaded with potential new meaning, and it's all happening in the mind of the spectator who wilfully (or accidentally)

¹ This is not to say that the performative cannot also be viewed semiotically. What I want to point to, however, is an effect of the performative that is not concerned with the communication/production of meaning.

theatricalises the situation by "misreading" it. MUDs provide numerous possibilities for such "misreadings", as written descriptions of objects provide little information compared to seeing objects in the real world. Additionally, most MUDs are already to a certain extent "theatrically framed" by explicitly being fictitious places, situated "elsewhere".

2.2 Performance

Performance is a much more complex concept than theatricality is. Because the term is used to describe phenomena within a lot of different disciplines, there exist several different definitions of it. Thus, in sociology performance could be connected to the representations of the self, while in linguistics it is the words that perform; when someone makes a promise, for instance. In theatre studies the word 'performance' can either generally refer to the staging of something, or specifically to the form 'performance art'. When used in relation to performance art, performance means 'real', the opposite of 'theatrical'. Generally used, and in the social sciences specifically, the meaning of the term seems to be closer to 'theatricality'.

Thus, while some disciplines will use a representational performance concept, others will insist that performance is non-representation. Theorists insisting on the non-representational aspects of performance will describe performance in terms of 'physical presence', 'non-representation', or 'non-referentiality'. This is common in performance art theory, where performance is viewed as the non-representational counterpart to representational theatre. (Féral 1992; Féral 1997/1982) Also the linguistic 'performatives' are described as non-referential, meaning that they do not refer to some external reality but constitute themselves acts. (Austin 1997/1955).

When playing a MUD, all we see is writing. Rooms, objects and beings are not *really* rooms, objects and beings; they are *written representations* of rooms, objects and beings. It could therefore seem paradoxical that I want to make use of a nonsemiotic performance concept in relation to MUDs. On the other hand, it is possible to define a 'real dimension' within the limits of the game, constituted by the game and isolated from the 'real world' taking place outside of it. Although affecting me - the real person in front of the screen - only emotionally, some effects of the game will severely affect my character. This will in turn affect the way I am able to proceed with the game.

3. MUD PERFORMANCE

3.1 Emotes and predefined verbs

In many MUDs there are two different ways to simulate character action. One way is to describe the action using the emote command, the other way is simply to type certain predefined verbs like 'smile', 'jump', 'greet player' or similar, which the program will then "translate" into happening.

Example:

1) emote

> emote smiles but doesn't look very happy.

Ragnhild smiles but doesn't look very happy.

2) predefined verb

> smile

Ragnhild smiles happily.

Using the emote command the player is free to define more or less exactly how she wants the action performed. Using the predefined verbs, she must rely on what is already defined in the program. There are, however, a lot of combinations possible, as most verbs can be combined with adverbs and/or modified in other ways.

Both emotes and predefined verbs are first and foremost used to communicate with the other players, that is, they are tools for communication and for playing role-playing games. Other kinds of games are not played against the other players, but against the program. Playing against the program, the player is "communicating" with preprogrammed objects and non-player characters. This kind of "communication" will be the focus of the rest of this paper. But first, let me present the theoretical framework for this part, the 'performatives' of J.L. Austin.

3.2 The performatives of J.L. Austin

In his book *How to do Things with Words* Austin defines a certain type of utterance that he calls 'performatives'. According to Austin these utterances differ from ordinary statements in that:

- A. they do not 'describe' or 'report' or constate anything at all, are not 'true' or 'false'; and
- B. the uttering of the sentence is, or is a part of, the doing of an action, which again would not *normally* be described as, or as 'just', saying something. (Austin 1997:5)

Examples of actions that are being performed this way, by the uttering of certain words under certain circumstances, could be marrying, promising or betting. By uttering 'I bet you sixpence it will rain tomorrow' (Austin's example), you make a bet. And this is important: you aren't just reporting it. Because you are making a bet and not merely reporting it, it makes no sense to ask whether this utterance is true or false. To be true or false a statement would have to refer to something other than itself, and as performatives aren't referring to anything, being acts and not reports, they cannot be true or false. They can however, be happily or unhappily performed, dependent on the circumstances under which they are being uttered. Not all circumstances would be appropriate for making a bet, for instance. If there are no other persons present, the words normally uttered to make a bet wouldn't have effect. The betting performance would thus be unsuccessful.

Performatives occur in MUDs too, of course. Representing fictional worlds more or less similar to the world we know, there are obviously instances of virtual marriages or of virtual betting. To take a more specific example, explaining how performatives 'work' in a MUD, consider the following: A highly ranked clan member of one of the TubMud clans will define another player as a member of her clan by typing the words 'welcome <name of player>'. Typing 'welcome <name of player>' is thus not to

make a constative or representational statement: it is to make the player a member, and the only possible way to do so too. Less obvious but not less performative are the commands with which the player moves around in the MUD. By typing 'n', or 'north', the player moves the character to the north. If there is an exit to the north, this performative will most likely be felicitous, if there isn't, or if the room to the north is open to wizards only, the character will not move north, and the performative will be infelicitous.

3.3 Theatrical emotes and performative verbs

While both the emotes and the predefined verbs from my first example will serve the communication and role-playing between the players first and foremost, some predefined verbs will work performatively too, dependent on the circumstances. The command 'welcome' is, for instance, in TubMud also functioning as a communications verb, when used by a character that has not reached the clan rank required to welcome new members. Thus, typed by the player of Edvard, an "ordinary" character, 'welcome <name of character>' will result in nothing but a message on the screens of the other player: 'Edvard welcomes you'. It's a friendly and communicative gesture, nothing more. When typed by the player of a clan leader, on the other hand, the action will have 'real' concequences: it will define the other character as a new member of the clan.

Character actions could thus be defined as more or less theatrical or performative dependent on the effects they are producing: emotional effects, affecting the player more than the character, or 'real' effects, affecting the character directly, or otherwise producing a change within the MUD world.

Emotes are normally not having any effect on the MUD environment. But as they are affecting the other players, in roleplaying environments they'll typically be effecting re-actions from the other players/characters². Thus, in a role-playing environment, emotes could be said to have performative effects. The same goes for the predefined verbs when these are used to communicate with the other players.

When used to communicate with objects and non-player characters, relevant predefined verbs should effect certain reactions from the object. When obvious approaches to interact with an object fail, MUDding is frustrating and boring. Extraordinary captivating room descriptions cannot, as some seem to believe, replace performativeness. Without performative interaction, the player will not only have to imagine the environment: she will have to *pretend* to be playing too.

4. QUESTING

4.1 Don Juan's promises

In her book *The Literary Speech Act*, Shoshana Felman is using Austin's performatives to analyse the Don Juan myth, based on her readings of the play *Don Juan* by Molière. Of performatives, it is the promise - or more precisely - the *broken* promise that defines and motivates the character Don Juan, and confuses and frustrates his antagonists. Felman describes the conflict between Don Juan and the others as a conflict between two orders; "the order of the act, and the order of meaning, the register of pleasure and the register of knowledge." (Felman 1983: 31)

A conflict between act and meaning is present in the activity of quest solving too. To do a quest is to search for the meaning of it. Having reached this meaning, the quest is solved. The paradox of questing is that as soon as meaning is reached, the quest stops functioning as *quest*. When meaning is found, the quest is *history*. It cannot be done again, as it is simply not the same experience to solve a puzzle quest for the second time.

In this, quests differ from ordinary, non-ergodic stories. The experience of re-reading a non-ergodic story isn't necessarily fundamentally different from the first time experience of reading it. This is because stories in general belong to the order of meaning, together with the constatives, and not to the order of the act. Quests, on the other hand, are basically performative: they belong first and foremost to the order of the act. As soon as they're solved, though, they turn into constatives. The reason quests can easily be confused with "stories" is that we are normally analysing the quest in retrospective, after we've already solved it. To ignore the performative aspect of quests this way is fundamentally to misjudge questing as a practice. Being acts before they are meaning, we must focus on the way quests *act* to understand the way they work.³

Felman writes: [...] Don Juan in fact does nothing but *promise the constative*. (Felman 1983: 35) This is exactly what quests are doing too, they are promising their solution, promising *meaning*. But as meaning is also the death of the quest, it is frequently breaking this promise, in order to prolong the questing experience.

4.2 Object reactions

Objects that the player finds in the quest area act as such *promises of meaning*. When the player finds a new object she will try to decode its significance in relation to the quest, to come closer to the quest's solution. To find the meaning of an object and determine its significance, the player must try out different ways of interacting with the object.

² By role-playing environments I mean MUDs that are explicitly devoted to 'role-playing'. Not all MUDs are equally serious about this role-playing component. Role-playing is done in interaction with other players/characters, and not in interaction with the program. Thus, role-playing a quest would require a game master to determine the performative effects that are otherwise determined by the program. Such performative effects will, however, not be further discussed in this paper.

³ Espen Aarseth has a similar point regarding the relationship between narrative and ergodic discourse in general: "Once realized, the ergodically produced sequence may be regarded and narratively reproduced as a story, but not one told for the player's benefit at the time of playing." (Aarseth 1999: 35)

Compare the following three programmed responses⁴ to a player's attempts of opening a box she finds in the MUD:

Example 1:

> open box What?

Example 2: > open box You can't open the box!

Example 3:

> open box

The box is locked.

> unlock box

You don't possess the right key.

Etc.

In the first example, there is no preprogrammed reaction to the attempt of opening the box, and thus the program responds "What?". The second example is a little less frustrating, with a reaction that at least appears to be intentional: "You can't open the box!". In the third example, however, the reaction is more engaging: "The box is locked" suggests that there is a key to the box, and "You don't posess the right key" suggests that a key is to be found, somewhere else.

Let's suppose that all three boxes are empty. They are not containing anything the player needs in order to solve the quest. In other words, they are nothing but stage properties, pure decorations, meant to serve the ambience of the quest environment. In such cases the player should perhaps not waste too much time trying to interact with them, as this isn't bringing her closer to the quest's solution in any case.

In the two first examples the player is also finished with the box after the first attempt of opening it. Ignoring the box she will be 'closer' to the quests solution, that is, its end. The reaction in the third example, though, may lead her on a detour seeking the missing key to be able to open the box. This detour will not affect the meaning of the quest at all, as it has no significance whatsoever regarding the quest's solution. The 'meaning' of this detour is to prolong the questing experience, to produce 'meaningful action'. Thus, to see the purpose of the detour, we can't focus on the quest's solution, we must focus on the quest as act.

Felman writes: "Paradoxically, the failure to carry out the promise makes it possible to begin it again: it is because the [...] promise is not kept that it can be renewed." (Felman 1983: 40)

Where the boxes from our first two examples reveal their insignificance on the player's first attempt of opening them, the third box shamelessly continues to promise. Until the player succeeds in finding the right key (or solves the quest), this (empty) box will continue to play a significant part in the game as a motivation for the player to keep on playing. When it comes to quest objects, thus, significance is not so much a question of representation, as it is question of performance. To the questing experience, it is more important how the object acts, than what it represents.

5. ACKNOWLEDGEMENTS

Many thanks to Michael Ruppel for proof reading. My research is funded by the SKIKT program of the Norwegian Research Council, and the University of Oslo.

6. REFERENCES

- Aarseth, Espen (1999): "Aporia and Epiphany in *Doom* and *The Speaking Clock*: Temporality in Ergodic Art", in Ryan, Marie-Laure (ed.) *Cyberspace Textuality*. Bloomington and Indianapolis: University of Indiana Press.
- Austin, J.L. (1997/1955): *How to do Things with Words*. Cambridge: Harvard UP.
- Carlson, Marvin (1996): *Performance. A critical introduction*. London: Routledge.
- Evreinoff, Nicolas (1927): *The Theatre in Life*, edited and translated by Alexander I. Nazaroff. London: George G. Harrap & Co.
- Felman, Shoshana (1983): *The Literary Speech Act. Don Juan with J.L. Austin, or Seduction in Two Languages*, translated by Catherine Porter. Ithaca: Cornell UP.
- Féral, Josette (1992): "What Is Left of Performance Art? Autopsy of a Function, Birth of a Genre", translated by Carolyn Tennessen, in *Discourse*, Spring, 14.2.
- Féral, Josette (1997/1982): "Performance and Theatricality: The Subject Demystified", translated by Terese Lyons, in Murray, Timothy (ed.) *Mimesis, Masochism, & Mime. The Politics of Theatricality in Contemporary French Thought.* Ann Arbor: The University of Michigan Press.
- Féral, Josette (1997/1988): "Teatralitet", translated by Eva Falck and Åste Redtrøen, in *3t*, 2.

TubMud telnet morgen.cs.tu-berlin.de 7680

⁴ It's a hypothetical example, constructed by me and not taken from any MUD. All responses are typical, however, they do occur frequently.

From "First Person Shooter" to Multi-User Knowledge Spaces

Mathias Fuchs Sibelius Academy, Helsinki and University of Applied Arts, Vienna

> Sebastianplatz 2/11, A-1030 Vienna +43 1 7148186

Mathias.Fuchs@uni-ak.ac.at

Sylvia Eckermann artist

Sebastianplatz 2/11, A-1030 Vienna +43 1 7148186 fuchs-eckermann@t0.or.at

ABSTRACT

We consider multi-user game engines such as Epic Megagames' UNREAL engine to be extremely useful tools for the design of knowledge spaces.

For a collaborative project with 10 Viennese museums¹ we developed a semantic matrix for a cross-disciplinary exhibition showing items from different collections (Sigmund Freud Museum, Jewish Museum, Museum of Natural History...). The content provided by these museums had to be made accessible and comprehensible to users of different ages, educational backgrounds and computer literacy. We developed a system of connotations amongst the objects, which then was translated into a spatial structure of rooms, corridors and places of different sizes, shapes, remotenesses or proximities. The viewer/listener of our knowledge space explores a semantic structure by navigating virtual spaces with the topics being contained in these rooms. The connecting architecture between these rooms resembles staircases, passages, elevators, hidden doors or portals, each of them referring to the nature of the connotation. Quite contrary to web-based databases and hypertext structures, the links therefore possess a quality of their own, carrying much more information than just "is connected with".

Keywords

Knowledge Space, Mnemosyne, Game Culture, First Person Shooter

¹ The opening of the project *Expositur – ein virtueller Wissensraum* was on May 2, 2001.

1. History

The concept for our computer-aided knowledge space is related to techniques of Mnemosyne, used by Greek "singers" (Simonides of Cheos) and philosophers as well as Renaissance scholars. [1] This form of mnemotechnique, called loci or place method, was widely used by orators to memorize complete speeches. The orator picked a building and learned every nook and cranny very intensely until he was able to move about the building in his memory. As a preparation for the speech a plethora of items of different complexity and amount of detail could be placed in the memorized rooms, e.g. a scale for justice etc. While delivering the speech the orator wandered from room to room and collected the hints while the speech unfolded.



Figure 1: Mnemotic system with abbey (Johannes Romberch: Congestorium Artificiose Memorie, Venice 1533)



Figure 2: Items used within the mnemotic system of Romberch

Obviously the urban structure of the mnemotic system helped in keeping a sense of orientation amongst the rather abstract items under investigation. That is why urban metaphors and architectural metaphors were so popular with medieval mnemotechniques.

In another respect our computer aided knowledge space also adopts techniques developed by Aby Warburg [2] for his research on the visual codes of Renaissance art. Warburg's scientific method consisted of connecting seemingly unrelated imagery to gain insight into visual similarities and connotations, which he called *Pathosformeln* [3][4]. In our knowledge space the multiple coding of meanings contained with the exhibited objects is made transparent by the spatial relation superimposed upon the objects. (A technical drawing of a prosthesis, e.g., is positioned close to Freud's *Prothesengott* quote and therefore connected to Freud's theory from "*Das Unbehagen in der Kultur*". The latter might lead to beautifully painted transportation vehicles from Pakistan which have been supplied to us by the Museum of Anthropology.)

2. Interface

To navigate the spaces of different content the users have to keep moving. They can walk, run, climb, jump, crouch, swim or fly according to the spatial situation. The Frankfurt based cultural scientist Manfred Fassler has mentioned in his recent publication that the etymological root of the German word for experience (*Erfahrung*) stems from *fahren*, i.e. "to move". We consider the process of actively exploring a quasi-spatial structure as the key mechanism for creating a semantic structure that is neither linear nor hierarchical. We consider the technology of a computer game as a helpful tool for the mediation of complex content. We also consider the freedom of the user to go his or her own way in the virtual environment as an important feature that allows for individually shaped relational networks inside a complex field of knowledge.

3. Developing a Structure for Knowledge Representation in Regards to a Concept of Spatial Organisation

Our approach of putting items containing multiply coded meaning into rooms which then are connected by other spatial structures poses a few questions which we tried to answer from a scientific standpoint as well as an artist's viewpoint:

1. How does the appearance of the rooms contribute to the meaning of the objects contained in these rooms?

2. How does the appearance of the connecting rooms contribute in signifying the relations between objects?

3. How does the appearance of the learning subject contribute to the learning process and the possible modes of acquiring knowledge?

4. How do modes of movement contribute to emphasize certain topics in the field of knowledge?

5. Is it essential for users to have a bird's eyes view of the semantic terrain they are investigating by cruising it?

6. What types of interaction can there be between multiple users exploring the same field of knowledge?



Figure 3: First person user in front of the Virtual Technical Museum (Virtueller Wissensraum by fuchs-eckermann)

Even though these questions might have been approached at a rather pragmatic level there is a semantic question lying underneath. If we are going to map semantic fields (*semantische Felder*) into fancy rooms, the signified (*Signifikate*) into topics, and the signifiers (*Signifikanten*) into 3D-objects or sound objects, does the game we are about to be immersed with still correspond to the semiotic structure we started with? If we suppose that we had a clear concept of a semantic structure, e.g. in the form of a Quillian model [5] or a Katz and Fodor structure [6], could we rediscover the model or structure in our game then? In other words: is there an isomorphism of the semantic structure

of the object under investigation and the spatial structure of the "level"?² We don't think so.

At least in the case of our approach there is a large amount of deliberate distortion of the semantic structure underway when we build the levels. We tend to concentrate on information which is feasible to be displayed in the game engine's range of possibilities. We also have to chop off branches of the semantic trees the technological platform we work with would not allow at present. Furthermore we often add connotations in the process of construction for the reason of subjective choice and mediaspecific deliberation. We do not intend to rebuild the museums or the structure of knowledge contained in the museums in electronic space. Our main goal is to find an artistic form of communicating topics which have traditionally been attributed to the museums. For that reason our work should be considered experimental.



Figure 4: Another user in front of the exhibition of prostheses (*Virtueller Wissensraum* by fuchs-eckermann)

The game engine we decided to use was not developed before 1999; therefore there are not many implementations of the kind of game we are interested in. Nevertheless, a few projects have been carried through which differ from the primitive look and feel of the original game. We are aware of the fact that our investigation is based on a very small sample of persons we could watch using the game. Our second knowledge space "Expositur" is at this moment only in a prototype state and still has to be considerably extended. Nonetheless there are a few hypothetical propositions we would like to make:

3.1 Appearance of Rooms

It seems of extreme importance for the creation of meaning how the environment the information is contained in, looks like and sounds: Within UNREAL we can create 3D-shapes for the rooms of a complexity of about 10,000 polygons. We recognized that the richness of the architectural forms decreases the stress a person feels confronted with, when navigating through this space. Former versions of UNREAL allowed for just a few hundred polygons and therefore favoured boxy levels creating a higher aggressiveness and a feeling of discomfort for certain users. On the other hand we recognized that very complex environments often create a feeling of "being lost" and of nausea. UNREAL Tournament at present allows using 8 sound slots at the same time, i.e. a number of 8 sound sources that may be set at any location in the rooms and can be heard simultaneously. The function of sounds accompanying the objects exhibited in the rooms can be analyzed in analogy to the function of sounds in movies:

3.1.1 Emotional Support

The sounds can deliberately be used to value objects as dangerous, hilarious, important, historically significant or other. We use sound in this respect to add ethical standpoints to objects we have to show, like armour, prostheses or extinct animals.

3.1.2 Additional Information about Visible Objects

We use recorded sound material to tell about the material qualities of objects.

3.1.3 Continuity Music

We recognized that turning off the background music of the rooms in a computer game results often in a much shorter playing time. Also the speed of the investigation, the restlessness and the carefulness of exploring a specific room can be manipulated via music played in the background.

3.1.4 Subtext

This is an important function for content creation of an ambiguous character or for the creation of content which can be interpreted in different ways. (In other words, this is important for any content.) For our game of knowledge we used the method of acoustical subtext in the room showing the collection of technical prostheses. The collection intends to throw a critical glance upon the dark sides of technological progress. In 1917 a workshop for prostheses was set up for workers having lost limbs when working on dangerous machinery. Of course in 1917 the number of persons being wounded in World War I was very high. The soundtrack accompanying the prostheses hints that capitalist production and warfare can result in the same sad results for the victims of either. This information, when put forward as a text, would sound quite banal and not lead to an intensive experience for most of the users. However, the soundscapes of heavy machinery mixed with heavy artillery and superimposed with composed rhythmical patterns does.

3.2 Appearance of Links

It has been mentioned earlier that we use staircases, passages, elevators, hidden doors and portals to signify the nature of the connotation, and in a second step: the nature of the relation between objects and topics. As Peter Kivy points out in his book "The Corded Shell" [7], there are "contour" features of visual objects and of acoustic objects, which are connoted with

² "Level" is game-talk for a subset of a complex field of knowledge. A level is in some way what one used to call a chapter in a book.

emotions. A rising line is usually considered to be positive or optimistic; a falling line to be negative, disappointing or dangerous. These features have to be carefully observed when laying out the spatial structure of the knowledge building. Also the size of passages can invite users to select the larger corridor for the further way instead of the smaller one. This is however not always the case. We recognized that frequent players often prefer the less obvious way instead of the more obvious one. This might have to do with an anticipation of tricks the game designer might have had in mind. It could also be that all users do not always enjoy the easiest way.

3.3 Appearance of the Learning Subject

UNREAL allows for the user to select "skins" from a set of stereotypes while running the game. We provide the same possibility for our knowledge space. After the avatar³ is visible when exploring the knowledge space, it influences the behaviour of the player. There are role models connected to certain "skins" which are very hard to predict. We found out that it is not necessarily the case that boys select male avatars and girls select females. It might be that the choice of an avatar has to do with what Denis Smalley calls levels of identification [8]. We suppose that certain users identify themselves strongly with the avatar whereas others find themselves outside the virtual space directing or leading a puppet through a simulated environment.

3.4 Modes of Movement

We consider the selection of a mode of movement by the player/user to be essential for the way the user structures his/her personal system of relevance. By freely selecting the speed, the user can declare topics as being relevant to him or of lesser importance. As opposed to e-learning systems based on video footage, the user of the knowledge space can find out his or her own pace in acquiring knowledge and can build his personal map of the field of knowledge. This is not always the case with computer aided systems. As Dave Campbell correctly observes, "software functionality like pre-defined paths of movement and pre-defined view-points of key spaces diminish a participant's ability to construct an accurate cognitive map of the space." [9] [10] We found out that different player types like to keep running all the time, or relax and just look around. UNREAL includes possibilities to force the user to move at certain points. The technical term for such objects is "movers", "teleporters" and "translocators". Movers are objects that can change location or move players when triggered. Movers might be used for situations resembling a daydream or a gentle shift of focus. Teleporters instantly move from one room to another. In terms of the knowledge represented by different objects, this means that you are taken over by a spontaneous idea or an abrupt change of mind. Translocators on the other hand are devices allowing the player to go from one point to another deliberately and quickly. In some way translocators can be compared to the rhetorical formula "Let's change the topic."



Figure 5: Running Avatar (Virtueller Wissensraum by fuchseckermann)

Another way of influencing the reception of content by changing the mode of movement of the player is changing the physical properties of the location or room the player is in. The designer of the knowledge space is not only able to change the gravity of the room but also to set the friction of the player on the ground. In this way certain topics can be presented on a ground which feels slippery or sticky. The player can be given the feeling of moving on thin ice or of being in a situation which makes it hard for him/her to move. It is obvious that this must have an effect on the care the player takes in acquiring the appropriate information. It probably also effects the level of satisfaction/frustration in gaining information.

3.5 Indexes, Survey

We are not sure whether a survey on the knowledge space is useful or desirable for the user. UNREAL provides the possibility to rise up into the sky (the process is called ghost mode) and take a look down upon the architectural structure of the knowledge space. The survey might increase the chance to go on once one is lost or caught in a labyrinth, at the same time it probably takes away the incentive for trying very hard to find out what one not yet knows.

3.6 Types of interaction

The modes of interaction UNREAL offers are not really useful for knowledge space design with multiple users. (They are with slight differences based on the imperative "Shoot them before they shoot you!") What we found most useful for real-time multiuser situations is the possibility for one knowledge seeker to show others their way by guiding them through the rooms. Another useful feature is the possibility to exchange messages via written or spoken word. This feature makes knowledge spaces of the kind described feasible for geographically spread networks of students, scholars and players. It is essential however to implement individual sets of voices and not to rely on the default voices provided by the manufacturer.

³ "Player Pawns" is what gamers call the avatars.

4. Implications

Multi-user games will play an important part in the field of elearning and the visualisation of knowledge. This fact is complained about by one side and advocated by the other. The discussion of the pros and cons of an audio-visual mode of knowledge representation versus a text-based one mirrors old discussions about the value, truth and appropriateness of images [11]. We would like to suggest however, that audio-visual virtual environments will gain importance during the next decade. The fact that the gaming industry gains ground compared to elder forms of edutainment like movies, video and printed text is an argument at least for not ignoring computer games. We think that these games - however simple they still might appear at present contain possibilities for knowledge spaces of a delicate nature - if they are thoughtfully conceived, carefully designed and joyfully experienced.

(Mathias Fuchs and Sylvia Eckermann,

e-mail: fuchs-eckermann@t0.or.at)

5. ACKNOWLEDGMENTS

Thanks to Manfred Fassler for critical discussions and to Renée Gadsden for proofreading the paper. The project has been carried out with the support of Austrian *Bundesministerium für Verkehr*, *Innovation und Technologie* and *Bundeskanzleramt Sektion für Kunstangelegenheiten*.

6. REFERENCES

[1] Yates, Frances A.: The Art of Memory. 1966.

- [2] Gombrich, Ernst H.: *Aby Warburg. Eine intellektuelle Biographie*, Frankfurt, 1981.
- [3] Roeck, Bernd: Psychohistorie im Zeichen Saturns. Aby Warburgs Denksystem und die moderne Kulturgeschichte In: Hardtwig, W. /Wehler, H.-U. (eds.), Kulturgeschichte Heute (Geschichte und Gesellschaft, Sonderheft 16), p. 231-254.
- [4] Wuttke, Dieter (ed.): Aby M. Warburg. Ausgewählte Schriften und Würdigungen, Baden-Baden, 1980.
- [5] Quillian, Ross M.: "Semantic Memory". In: Minsky, Marvin (ed.): Semantic Information Processing. MIT Press, Cambridge, 1968.
- [6] Katz, J. J. and Fodor, J. A. (eds.): The Structure of Language. Prentice-Hall, Englewood Cliffs, 1964.
- [7] Kivy, Peter: The Corded Shell. Princeton University Press, Princeton N. J., 1980.
- [8] Smalley, Denis: Spectromorphology: Explaining Sound-Shapes. In: Organised Sound. Vol. 2(2), Cambridge University Press, 1997.
- [9] Campbell, Dace: Design in Virtual Environments Using Architectural Metaphor, New York 1995, pp.64.
- [10] Daniela Bertol: Designing Digital Space. An Architect's Guide to Virtual Reality, New York 1995, pp. 88.
- [11] Müller-Funk, Wolfgang: Konstruktion des Erinnerns: Transitorische Turbulenzen 3. Mnemo-Topologien: Die Abwesenheit Gottes. Kunstforum Cologne July - September 1994, pp. 216.

The text you write must desire me: fonts as the penultimate interactive artform, second only to sex

Paul Chan National Philistine 305 W. 28th St, SUITE 19C New York City, New York 10001 www.nationalphilistine.com info@nationalphilistine.com

ABSTRACT

In this paper I discuss and present *Alternumerics*, a collection of computer fonts that reinterpret text.

1. INTRODUCTION

explores the intimate relationship between language and interactivity by transforming the simple computer font into an art form that explores the fissure between what we write and

By replacing the individual letters and numbers (known as alphanumerics) with textual and graphic fragments that signify what is typed in radically different ways, *Alternumerics* transforms the act of typing into a digital performance and any computer connected to a standard printer into an interactive artmaking

Alternumeric fonts are Macintosh and Windows compatible and work with any application that uses fonts. Each font is accompanied by an art-piece that uses the font to explore the relationship between what is typed, what is translated—and fundamentally—what is communicated when we use language to describe the pleasures of utopia (see Figure 1), the slipperiness of the self (see Figure 2), the friction of desire (see Figure 3), and the poetry of silence (see Figure 4).

Alternumerics has been exhibited in new media festivals and exhibitions worldwide. The fonts can be downloaded on-line at:

www.nationalphilistine.com/alternumerics/

2. HISTORY

I don't remember exactly why I began to mutate fonts into forms that both reduce and expand its signifying possibilities. It wasn't as if language had stopped working for me. I could still signify love and longing and malice and the infinite space of the future through the given set of alphanumerics on my keyboard: I could still write. But I wanted more. I got greedy. I wanted language to only work for me and no one else.

3. PRACTICE

It is relatively easy making fonts. Unlike other new media art practices, the technologies used to make fonts have remained stable. The politics of perpetual obsolescence in technology forces most new media art into a state of permanent retardation. Technology should never dictate the form; it can only dictate

The field that fonts play in is expansive and intimate. It is loaded into your computer on a systems level, so any application that uses different fonts can play. Word processing applications become a canvas for the lines and shapes and markings that make up the materiality of fonts. Database software is transformed into a De Sadean regulator of pies charts and graphs that regulate signifying perversities. Did I mention how small fonts are? Fonts are invariably under 100K and virtually any computer can work Ubiquitous. Viral.

4. THOUGHTS

It is useful for me to think about my work with fonts as systems that bleed. What I have essentially done is reduce the material possibilities of fonts to signify the immaterial by making the material more specific, more historical, less universal, and more accountable to me. And like any system that reduces the world it is inherently tragic. Think Diderot's Encyclopedia. Think

The tragedy and promise of my fonts come from the residue that is left when I reduce the letters and numbers and symbols on the computer keyboard into something other than the signifying practice of a time and a space that neither cares for me nor thinks about me. The residue is a reminder of the unsayable in my font, or any font for that matter. The residue is that which cannot be reduced into language. It is body. It is blood. It is desire that

I also sometimes think of my fonts as ghosts that take over your body so that they may enunciate a history and a longing that is

Figure 1







The future must be sweet

Self portrait as a font V.2

Figure 3





Figure 4

Sexual healing / Shift for harassment

Blurry but not blind

THE FUTURE MUST BE SWEET (2000)

Utopian Socialist Charles Fourier believed the world should be organized around our pleasures. Politics become the body that regulates and maximizes our desires by ensuring every one equal access to affection, justice, and exquisite food. This font reinterprets Fourier's philosophy into a textual-graphic

system and gives form to the unique connections Fourier made between radical politics and utopian desires. Different relationships between the letters (and words) develop based on simple changes in word processing: point size, page width,

CALANTES	11771 P 1186		<u>_</u>			S BRUULANT LOUT
amorous corporations	—but gingerly	a longing	REVOLUTION domestic destiny	desire	friendship	the social compass
WIVES	SWEET CREAMS	PHILANTHROPY	FINE WEATHER	A SOCIAL MOVEMENT	£	BRIGHT PASTRIES
а	D	C	a	е	T	g
×.	human passion	SPLENDID TOPIARIES	DISTRIBUTE IT	FATTY DUCK	BLISS	LIBERATE
1ncoherent household	A PROMISE social mechanism	-a phalanx RADIANTLY WEALTHY	luxury {external} PARIS IN WINTER	love	{ internal } { hó tai	southern fluidfemale
h	i	j	k	I	m	n
AN INSUPPORTABLE EMPTINESS	<u> </u>	GIRAFFE	usefulness	BODILY WEAKNESS	MONEY	
go on	paternity	truth	pleasure		unescapable EXCESS AFFECTION	barbarism РНПОВОРНУ
0	р	r	S	t	u	v
progressive	HYPERBOLA	castrate northern		(altornationa)	(buanni as)	Scapitalists]
series MANGOS		fluid MALE	HAPPINESS CHAOS	to this ugliness	{bunnies}—	{capitatists}=
W	х	У	Z	A	В	С
{deers}—	{endives}—	{future}—	{great}	{hope}	{industrialists}—	{justice}—
D	Е	F	G	Н	I	J
{kinship}—	{lovers}—	{marmalade}-	{necessity}–	{opulence}—	{philosophers}—	{quiche}—
К	L	М	Ν	0	Р	Q
{ roasts } preferably pork	{socialists}—	{table wine}-	{utopia}—	{violence}—	{wafers}—	{unknown}–
R	S	Т	U	V	W	Х
{yams} —	$\left\{ \overset{a \text{ sout}}{\mathcal{Z}00} ight\}$ ———	——a sorry state	HOPE	-{civilization}	{fiction}—	włopia is a stato of society where marx would no longer critique fourier – warms
Y	Z		connoun reason	?	+	~

SELF PORTRAIT AS A FONT V.2 (2001)

What is it like to write like me? You don't even know me. Lowercase letters are phrases I think I say in conversations. Uppercase letters are parenthetical comments based on what I think I say in conversations and common mistakes I make when writing. Numbers are names of friends, family, and former lovers. Symbols are incidental words I use to feign interest, confusion, or indifference. And yes, it is my

I	Don'T REMEMBER	WON'T BE SURE	NEVER GOT IT	I	WILL FORGET YOU	BARELY COMPREHEND
а	b	С	d	е	f	g
DION'T MEAN THAT	BECAUSE I	DON'T FOLCOW	WILC FORGET IT	IHAGINE	PROBABLY HISREAD IT	APOLOGIZE
h	i	j	k	I	m	n
SOHETIMES I	WANT TO EAT	DON'T LIKE IT	PLAIN FORGOT	NEVER KNOW	OR MAYBE YOU	DON'T UNDERSTAND
0	р	r	S	t	u	V
REALIZE THE GRAVIT	Y MAY NEVER FNOW	Don't MEAN MUCH	REMEMBER IT DIFFERENTLY	Margan	(MOST LIKELY)	(FUNNY)
w	х	У	Z	А	В	С
(HONESTLY)	(NV6)24000	(IT'S TOUGH, I KNOW)	(I'M NOT KIDDING)	(ITS A BIT SAD)	- ATMAN	(WHATEVER THAT MEANS)
D	Е	F	G	Н	I	J
(DO I REALLY MEAN THAT?	") (I'M SORRY)	(MARX SAID IT, NOT ME)	(NO DTHER WAY I THINK)	(But new rhad Statemeters)	(and that's that)	(ALL <u>UERY</u> CONFUSING)
К	L	М	Ν	0	Р	Q
(HOPE THIS HELPS)	(IT'S NOT THAT FUNNY)	(IDON'T GET IT)	200 For the	(UOLTAIRE WOULD KNOW)	(IT'S MEANINGLESS) (ITS NOTHING)
R	S	Т	U	V	W	Х
(yes? maybe?)	(Is that All ?)	- AMINOR PAUSE	-SILENCE-	CAN'TTHINK STRAIGHT.	SLEE P	WITH ME
Y	Z	,		?	()

SEXUAL HEALING / SHIFT FOR HARASSMENT (2000)

Lowercase letters are phrases taken from popular love songs of the 70's, 80's, and 90's. Uppercase letters are phrases taken from transcripts of sexual harassment cases in the United States from the 70's, 80's, and 90's. Numbers and symbols

are words that heighten the tension between the play of the uppercase and lowercase letters as they shift between the voice of pleasure and the voice of violence.

so hot	sweet thang	touch me	faster	(you)	hold me	tonight
а	b	с	d	е	f	g
oh god	(me)	don't go	freak me	the pleasure	love me more	feels nice
h	i	j	k	Ι	m	n
oh	you complete me	I want you	don't stop	oh girl	(he)	oh beautiful
0	p	r	S	t	u	V
I feel it	let's do it	so silky	so much love	stop	I mean it	get off me
w	x	у	Z	А	В	С
please stop	don't	it hurts	I'm bleeding	that's enough	oh god stop	the pain
D	E	F	G	Н	I	J
let go	don't do this	hands off	grow up	help me	please don't	you're hurting me
К	L	М	Ν	0	Р	Q
help me	rape	back off	(he)	it's not funny	someone help	I'm begging you
R	S	т	U	V	W	Х
can't breathe	not happening	(screams)	(silence)	(in a whisper)	no more	(with fanfare)
Y	Z			?	:	١

BLURRY BUT NOT BLIND —AFTER MALLARMÉ (2001)

"The imperfection of languages consists in their plurality, the supreme one is lacking: thinking is writing without accessories or even whispering, the immortal word still remains silent; the diversity of idioms on earth prevents everybody from uttering the words which otherwise, at one single stroke, would materialize as truth." —Stephen Mallarmé. This font formalizes Mallarmé's idea that silence is the true universal language. Lowercase letters are empty kerning spaces of varying lengths. Uppercase letters are empty kerning spaces and typographic symbols inspired by Mallarmé.

L				U		
white space	white space	white space	white space	white space	white space	white space
а	d	С	a	е	1	g
		white space	white space		white snace	
h	i	i	k		m	n
	·	,	i i i i i i i i i i i i i i i i i i i	·		
Ц			Ц	U		
white space	white space	white space	white space	white space	white space	white space
0	р	r	S	t	u	V
					some breezv	
white space	white space	white space	white space	white space	y	and (see)
W	x	v	Z	A	В	С
		,				
warming	()	+ some	L	L	()	L
			white space	white space		white space
D	E	F	G	Н	I	J
				Like		
white space	()		()	a g(host)	here	white space
K	L	M	N	0	Р	Q
	-			-	·	-
L	come	L1		L		?
white space		white space	white space	white space	white space	
R	S	Т	U	V	W	Х
	break the					
remember	day +	white space	white space	レ white space	*	white space
Y	Z			?	0	=

(IDON'T GET IT) DIDN'T MEAN THAT I ... WWW POLOGIZE NEVER GOT IT - SILENCE -

> The End.

Bridging the Semantic Gap in Content Management Systems: Computational Media Aesthetics

Chitra Dorai IBM T. J. Watson Research Center P.O. Box 704, Yorktown Heights New York 10598 USA dorai@watson.ibm.com

ABSTRACT

With the explosion of online media and media-based services, a key challenge in the area of media management is automation of content annotation, indexing, and organization for efficient access, search, retrieval, and browsing applications. One of the major failings of current media annotation systems is the semantic gap which refers to the discontinuity between the simplicity of features or content descriptions that can be currently computed automatically and the richness of semantics in user queries posed for media search and retrieval. This paper proposes an approach that targets at bridging the semantic gap and building innovative content annotation and navigation services. The approach is founded upon an understanding of media elements and their role in synthesis and manipulation of program content with a systematic study of media productions. It proposes a framework for computational understanding of the dynamic nature of the narrative structure and techniques via analysis of the integration and sequencing of audio/visual elements. The resulting system will lead to automatic content organization and interpretation that provides high level and high quality content descriptions to aid in search, retrieval, and browsing and also to objective and consistent distillation of the common features of successful audio-visual strategies.

1. INTRODUCTION

While issues of media archival as well as delivery on the Internet and corporate intranets are adequately addressed by improved compression standards, faster networks, and advances made in storage and streaming technologies, the challenges of automating media annotation, content indexing, segmentation, and organization for search, retrieval, and browsing applications are still being tackled. Automatic content indexing and annotation is a growing area of research in media computing, and a recent survey paper summarizes the state of art and identifies the key challenges [14]. The failing of current systems is that while "the user seeks semantic similarity, the database can only provide similarity on data processing". The authors define the semantic gap as the "lack of coincidence between the information that one can extract from the visual data and

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands. Svetha Venkatesh Department of Computer Science Curtin University of Technology GPO Box U1987, Perth, 6845, W. Australia svetha@cs.curtin.edu.au

the interpretation that the same data has for a user in a given situation" [14]. The discontinuity between the simplicity of features or content descriptions that can be currently computed automatically and the richness of semantics in user queries posed for media search and retrieval makes user acceptance and adoption of automated content annotation systems very difficult. The authors of the survey conclude that "bridging the semantic gap between the simplicity of available visual features and the richness of user semantics" is the key issue in building effective content management systems.

To address this issue, we depart from existing approaches to deriving video content descriptions [8, 12, 11, 13]. Motivated and directed by video production principles, we propose an approach that goes beyond representing what is being directly shown in a video or a movie, and aims to understand the semantics of the content portrayed and to harness the emotional, visual appeal of the content seen. It focuses on deriving a computational scheme to analyze and understand the content of video and its form. Accepted rules and techniques in video production are used by directors worldwide to solve problems presented by the task of transforming a story from a written script to a captivating narration [5]. These rules, termed as film grammar in the movie domain, refer to repeated use of certain objects, visual imagery, and patterns in many films to instantly invoke a specific cinematic experience to the viewers [16, 15]. The rules and icons serve as shorthand for compressing story information, characters, and themes into known familiar formulae, often becoming the elements of a genre production. They constitute a style or form of artistic expression that is characteristic of content portrayed, and can be considered to be almost idiomatic in the language of any program composer or director. Production rules are found more in history of use, than in an abstract predefined set of regulations, and elucidate on ways in which basic visual and aural elements can be synthesized into larger structures.

We hypothesize that the employment of these tacitly followed rules in any genre not only can be understood and derived automatically with a systematic study of media productions, but also be exploited in characterizing what is happening in a video for high-level video/film abstraction in an algorithmic framework we term, *Computational Media Aesthetics*. The framework allows for a computational understanding of the dynamic nature of the narrative structure and techniques via analysis of the integration and sequencing of audio/visual elements, and is targeted at bridging the semantic gap and building effective content management systems at higher levels of abstraction and meaning. Further, it puts video/film analysis on a sound footing resting on principles and practices from video/film production rather than on ad hoc schemes. While earlier work [7, 9] using film grammar has focused on content generation, synthesis of video presentations, and virtual worlds, our emphasis is on characterizing, describing, and structuring of produced videos for media search, segment location and navigation services.

The layout of this paper is as follows: in Section 2 we define and discuss the framework, Computational Media Aesthetics. The lessons we learned in the use of this framework to extract two higher order semantic constructs from film are discussed in Section 3. Our conclusions follow in Section 4.

2. COMPUTATIONAL MEDIA AESTHET-ICS

Zettl defines Media Aesthetics as a study and analysis of media elements such as lighting, motion, colour and sound both by themselves and their roles in synthesizing effective productions [16]. We define *computational media aesthetics* as the algorithmic study of a number of image and aural elements in media and the computational analysis of the principles that have emerged underlying their use and manipulation, individually or jointly, in the creative art of clarifying, intensifying, and interpreting some event for the audience.

What does this new framework entail? By focusing on the emotional and visual appeal of the content, it attempts to uncover the semantic and semiotic information by a study of the relations between the cinematic elements and narrative form. It enables distilling techniques and criteria to create efficient, effective and predictable messages in media communications, and to provide a handle on interpreting and evaluating relative communication effectiveness of media elements through a knowledge of film codes that mediate perception, appreciation and rejection.

Our computational framework analyses and understands the content of video and its form. This approach, undergirded by the broad rules and conventions of content creation, uses the production knowledge to elucidate the relationships between the many ways in which basic visual and aural elements are manipulated in video and their intended meaning and perceived impact on content users. Our computational scheme analyzes videos to understand the film grammar, in particular and uses the set of rules that are commonly followed during the narration of a story, to assist us in deriving the annotation or description of video contents effectively. A system built using this principled approach where videos are analyzed guided by the tenets of film grammar will be effective in providing high-level concept oriented media descriptions that can function across many contexts and in enhancing the quality and richness of descriptions derived. We propose a two-tiered framework: Primitive feature extraction and a complex higher order semantic construct extraction stage (See Figure 1).

2.1 Primitive Feature Extraction

In our approach, first, like those of other researchers, simple features such as colour, motion, editing effects, sound signal energy, etc are extracted. Given a video of a movie, news program, a class or a training program, shot segmentation is carried out to partition the video into atomic units for further processing. Based on the low level visual and aural attributes, various shot-based attributes are computed: shot duration, average number of shots per unit time, its variance, shot colour features, average shot motion and variance, changes in perceived visual motion, shot audio energy level, etc. These features can be directly computed from frame or shot processing.

2.2 Higher Order Semantic Construct Extraction

What sets the framework apart from other schemes is this stage. Here, we extract complex constructs, or expressive elements that expose the underlying semantic information embedded in the media production. The extraction of increasingly complex features from a hierarchical integration of underlying primitives is a commonly followed approach. But the key difference is this framework of analysis based on production knowledge, that is, to both define what to extract, and how to extract these constructs we seek guidance from film grammar. We do so because directors create and manipulate expressive elements related to some aspect of visual or emotional appeal in particular ways to have maximum impact. With movies for example, we draw attention to the film creation process, and argue that to interpret the data one must see it through the filmmaker's eye. Film grammar is the portal that gives us insight into the film creation process. It can tell us not only what expressive elements a director manipulates, but also how she does it, why, and what the intended impact is. Thus, complex constructs are both defined and extracted only if media production knowledge tells us that it is an element that the director crafts or manipulates intentionally. These elements by their derivation and study result in crafting human-friendly content descriptions since they directly impact viewers' engagement with the content portrayed.

These complex constructs typically cannot be extracted directly from the shots, like the primitive features, but are built upon them. Many cinematic techniques, operating on many media elements such as color, camera movements, and sound contribute to the creation of an expressive element, and therefore an integrated analysis of multiple low level features across sequences of shots becomes essential. We are aware that video production grammar may indeed lead us to some expressive elements that do not easily translate into algorithms. However, several expressive elements are based upon manipulation of physical elements such as objects and cameras, and therefore can be defined in terms of their operations and extracted. Pace, rhythm, and tone are examples of some of the first set of complex constructs that can be examined and studied.

3. CLOSE ENCOUNTERS WITH EXPRES-SIVE ELEMENTS IN FILM

To explore the feasibility of our framework and test its efficacy we have investigated the following expressive elements in film: pace and rhythm. We outline in each case what we learned from film grammar, and then present our observations related to the extraction of these two expressive elements.

3.1 Pace, According to Film Grammar

Tempo or Pace is often used interchangeably in film appreciation, and refers to the "rate of performance or delivery". Zettl [16] makes a distinction in defining pace as the perceived speed and tempo as the perceived duration. Thus tempo/pace is a reflection of both the speed and time of the underlying events being portrayed and affects the overall sense of time of a movie. Tempo is crafted and manipulated in different ways. One technique is the montage that allows a director to manipulate the shot lengths used in the creation of a scene, thus deliberately controlling the speed at which a viewer's



Figure 1: The Computational Media Aesthetics framework.

attention is directed. Another means by which a viewer's perception of speed can be manipulated is through controlling object and camera motion. Fast motion gives us the feeling of fast events, while no or little motion has the opposite effect on our perception of pace. Film audio is a third factor that increases or decreases our sense of the performance delivery. There may be other more subtle factors besides the story itself, but we argue that one can construct a computable and powerful expressive element, pace, that reasonably captures the flow of time in a movie based on the underlying primitives of shot length and motion.

3.2 Algorithms and Implications

Based on our understanding from film grammar that pace is primarily affected by shot length and pace, we define P(n), a continuous valued pace function as

$$\boldsymbol{P}(\boldsymbol{n}) = \alpha(W(s(n))) + \frac{\beta(m(n) - \mu_m)}{\sigma_m}.$$
 (1)

where s refers to shot length in frames, m to motion magnitude, μ_m and σ_m , to the mean and standard deviation of motion respectively and n to shot number. The weights α and β , refer to relative

weights that affect the extent to which shot length and motion contribute to pace. Without any *a priori* knowledge, they can be given values of 1, effectively assuming that both shot length and motion contribute equally to the perception of pace for a given film. It is possible, however, that under certain circumstances one or the other of these two impact more heavily on the audience perception of time, depending on the movie genre or a director's style. W(s(n))is an overall two part shot length normalizing scheme, having the property of being more sensitive near the median shot length, but slows in gradient as shot length increases into the "longer" range.

We have examined this function for several motion pictures including the Titanic, Colour Purple, Lethal Weapon 2 and The Matrix, and can make the following observations:

• The ebb and flow of the pace function **P**(**n**) delineates the dramatic flow of the content and concurs very reliably with a qualitative assessment of movie tempo. The pace function paves way for characterizing movie content in many interesting ways, such as to organize content in terms of their dramatic import, to quantitatively compare and summarize

movie sections based on their tempo measure, or even to institute different policies to reduce our graduated tempo measure to labels, if desired to textually annotate content.

- Significant pace changes occur at the boundary of story sections, and often coincide with events of high dramatic import. Relative pace changes across the movie can be determined from the pace function. We extracted the edges of the pace function with a multi-resolution analysis, and found that large pace transitions coincide with what we term as story transitions. Smaller pace transitions coincide with what we term events, and these are differentiated from story transitions in that they occur more locally and contained in time. For example, the pace transition between the first and third class party scenes in the Titanic is a large one. The first class party scene is shown as sedate and slow. The third class party is vivacious and merry. As they happen juxtaposed, the pace difference between these two story sections is big and well captured. On the other hand, when Jack grabs Rose as she attempts to fall off the ship, there is a brief flurry of action, leading to a smaller pace transition which we label as a localized event. Our analysis has revealed that we can detect pace edges efficiently and that both story sections and events can be located (for further detail see [3, 2, 1]).
- Pace, therefore, is found to be high-level and fundamental (applicable to multiple contexts), yet manifest in a way to be computationally tractable. It offers pointers to automatically organizing videos into higher-level segments characterized by their dramatic development and to their high-level descriptions.

Results from The Matrix, one of a number of movies analyzed are presented here for the purpose of demonstration. Figure 2 shows the pace plot of a section of the movie with located edges indicated for two of the 4 Σ/τ combinations used in edge detection, and Table 1 matches each automatically discovered edge to a brief description of the *story section* bounded by, or the dramatic *event* coinciding with the discovered edges.

 Table 1: Labelled story sections and events identified from tempo changes in The Matrix (cf. Figure 2).

Story Element detected (high thresh)				
А	Morphius in captivity			
В	The rescue of Morhpius			
С	Neo finally confronts the matrix			
D	Neo is the 'one'			
Eve	ent detected (low thresh)			
а	Neo and Trinity face the soldiers			
b	Trinity shoots the agent			
с	The cavalry (helicopter) arrives to rescue Morphius			
d	Morphius leaps to safety			
e	The escape			
f	An agent appears, Trinity escapes			
g	Neo finally faces an agent			
h	and wins			
i	Neo is chased			
j	¡False positive¿			
k	Neo is killed?			
1	No, he is alive, the climax			



Figure 2: Results of edge detection on pace flow and corresponding story sections and events from the Matrix.

3.3 Rhythm in Film

Film rhythm is another complex narrative concept used to endow structure and form to film. Mitry defines it as an "organization of time" [10]. Of the many, often elusive cinematic devices contributing to film rhythm, Bordwell and Thompson [6] state that "frame mobility involves time as well as space, and film makers have realized that our sense of duration and rhythm is affected by the mobile frame". They list camera position/movement, sound rhythm, and editing as constituent elements of rhythm. Further they label resulting rhythms types in higher level terms by stating that a "camera motion can be fluid, staccato, hesitant and so on".

Thus, because a film is structured in time with editing, it manifests a natural beat, and has an intrinsic rhythm. To find this rhythm, one must examine a neighborhood of shots. In addition, since both shot length and motion contribute to rhythm, one can examine the rhythm that arises individually and jointly from these contributing elements. Since shot length and motion are computable, motion rhythm and editing rhythm are likewise derivable from them.

3.4 Computing Rhythm

We have chosen to study two main contributing factors to rhythm, namely the shot length and motion. We examine each separately (see [4] for details).

Shot rhythm: Since shot lengths can be crafted in many ways to produce almost an infinite variety of beats, we restricted the initial shot rhythm classes to the following:

- Metrical: in which a series of shots have nearly the same shot length.
- Attack: in which a series of shots have steadily decreasing shot length.
- Decay: corresponds to monotonically increasing shot length.
- Free: None of the above, and therefore in these sections of the movie, shot rhythm is secondary instead of being the driving factor behind shot placement and length.

We implemented an algorithm to extract and classify shot neighbourhoods with their shot rhythm label. We have experimented with both the neighbourhood size and the tolerances required to classify movie sections robustly. We make the following observations: (a) Attack is often used to lead into a dramatic event. A rising shot rate captures and leads the viewer to the events of high dramatic import that are about to follow. (b) It does not appear that decay is crafted to the same extent as attack, but it has the opposite effect to attack in perception. (c) Even if the identification of attack, metric and decay sections does not tell us about the exact content, they act as clear signs that the director is creating or reinforcing *something* by their use. (d) A precision of about 70-75% was registered in automatically classifying sections with these shot rhythm labels.

Motion Rhythm: Film grammar guides us to divide the classes of motion rhythm as

- No Motion,
- Fluid: Where the motion is smoothly changing or constant, and

• Staccato: where the motion contains abrupt speed changes.

A shot motion behaviour classifier was implemented and the results analyzed across a range of movies. Our observations from these experiments are: (a) The classifier in labelling the motion rhythm classes achieved an accuracy of 70-80%. (b) The most interesting results from examining motion rhythms arose from the degree of common scene content found in different rhythm categories. (c) Neighbourhoods with predominantly fluid rhythm generally correspond to long and establishing shots. (d) Staccato motion neighbourhoods are very taxing and used sparingly, particularly in extremely violent scenes. (e) A mixture of Fluid and Staccato correspond to scenes where a dual perspective is being shown, for example, action sequences.

3.5 Inferring Content Semantics

What inferences can we draw by a joint analysis of both shot and motion rhythms? How do they interact? What clues do they provide for automatically structuring content? To answer these questions, we studied both shot and motion rhythms across several movies. Our main conclusions are the following:

It is not possible to make a deduction on the actual content of a scene from its rhythm. However, one can reliably infer that when rhythm changes, something has changed in the scene.

We can also conclude something about why the content has changed. A shot device (either metric, attack or decay rhythm) accompanied by a motion rhythm change generally corresponds to a scene change being precipitated by an event, or to the reinforcement of change. A free shot rhythm accompanied by a motion rhythm change generally implies a change to a different locale, scene or sequence abruptly. A shot device (metric, attack or decay) with no accompanied motion change generally implies the same film setting or a result of an event.

Rhythm, too therefore, like pace, is computable. It leads us to understand where scene content has changed, and hypothesize as to why this could have occurred. This change and its reasoning is at a much higher semantic level than could have been done reliably by merely examining the change of a low level primitive like shot length.

These experiments present some of our initial findings. We are currently gathering experimental data to support distinctive occurrences of shot and motion rhythm classes across genres, in order to understand whether we can index the content at a finer resolution.

4. CONCLUSIONS

We propose to bridge the semantic gap by using media production understanding to guide media analysis. We believe that it is this understanding that will enable us to both formulate and extract the correct semantic entities, and enable us to structure video and film automatically. These semantics can lead to the development of shared vocabularies for structuring video and film, and serve as the foundation for media description interfaces. While we have drawn upon film grammar to derive the expressive elements in film, we believe that such an approach will work in other video domains. News, Sitcoms, Sport etc. all have more or less complex grammars that may be used to capture their crafted structure. There is structure regardless of particular media context but there may not be homogeneity, and therefore it helps to be guided by production
knowledge in media computing.

We are building upon our initial results to eventually develop videowatching agents or systems that can make human-like judgements about images and sounds in film, and will lead to finely tuned, personalized screening and practical filtering tools. The software models will enable technologies that can emulate human perceptual capabilities on a host of difficult tasks such as parsing movie into sections of interest, making inferences about movie semantics and about the perceptual effectiveness of the messages contained, and estimating the similarity of two productions.

5. ACKNOWLEDGEMENTS

We thank Brett Adams for being a willing and helpful partner in our close encounters with the expressive elements explored so far.

6. **REFERENCES**

- B. Adams, C. Dorai, and S. Venkatesh. Role of shot length in characterizing tempo and dramatic story sections in motion pictures. In *IEEE Pacific Rim Conference on Multimedia* 2000, pages 54–57, Sydney, Australia, December 2000.
- [2] B. Adams, C. Dorai, and S. Venkatesh. Study of shot length and motion as contributing factors to movie tempo. In 8th ACM International Conference on Multimedia, pages 353–355, Los Angeles, California, November 2000.
- [3] B. Adams, C. Dorai, and S. Venkatesh. Towards automatic extraction of expressive elements from motion pictures: Tempo. In *IEEE International Conference on Multimedia and Expo*, volume II, pages 641–645, New York City, USA, July 2000.
- [4] B. Adams, C. Dorai, and S. Venkatesh. Automated film rhythm extraction for scene analysis. In *IEEE International Conference on Multimedia and Expo*, Tokyo, Japan, August 2001, To appear.
- [5] D. Arijon. Grammar of the film language. Silman-James Press, 1976.
- [6] D. Bordwell and K. Thompson. Film Art, 5th Ed. McGraw-Hill, 1997.
- [7] M. Davis. Knowledge representation for video. In Proceedings of the Twelfth National Conference on Artificial Intelligence, pages 120–127, 1994.
- [8] J. Huang, Z. Liu, and Y. Wang. Integration of audio and visual information for contend-based video segmentation. In *Proc. IEEE International Conference on Image Processing*. IEEE, 1998.
- [9] C. Lindley. A computational semiotic framework for interactive cinematic virtual worlds. In *Workshop on Computational Semiotics for New Media*, Guildford, Surrey, UK, 2000.
- [10] J. Mitry. *The Aesthetics and Psychology of the Cinema*. The Athlone Press, London, 1998.
- [11] J. Nam, M. Alghoniemy, and A. Tewfik. Audio visual content based violent scene characterization. In *Proc. IEEE International Conference on Image Processing*. IEEE, 1998.

- [12] V. Pavlovic. Multimodal tracking and classification of audio visual features. In *Proc. IEEE International Conference on Image Processing*. IEEE, 1998.
- [13] C. Saraceno and R. Leonardi. Identification of story units in audio visual sequences by joint audio and video processing. In *International Conference on Image Processing*. IEEE, 1998.
- [14] A. Smeulders, M. Worring, S. Santini, and A. Gupta. Content based image retrieval at the end of the early years. *pami*, 22(12):1349–1380, 2000.
- [15] T. Sobchack and V. Sobchack. *An introduction to film*. Scot, Foresman and Company, 1987.
- [16] H. Zettl. Sight, Sound, Motion: Applied Media Aesthetics. 3rd Edition, Wadsworth Pub Company, 1999.

Introducing Interrelations Of Madness & Virtuality

Steffen P. Walz Institute for Social Anthropology University of Tuebingen 72070 Tuebingen, Germany +49 7071 29 72374

spw@playbe.com

Isabel Zundel playbe Kornbergstr. 36 links 70176 Stuttgart, Germany +49 711 28417 35

izun@playbe.com

ABSTRACT

This paper approximates interrelations between representations of madness, and virtuality in Immersive Virtual Environments semiotically. Thereby, social and sensual performances indicating coded acts of practice are mapped, and investigated from several perspectives. The authors give an introduction to these 'altered states', and suggest that it is possible to learn from codes and mappings of madness in order to create virtual environments, and learn from codes of virtual environments in order to understand representations of madness.

General Terms

Design, Experimentation, Human Factors, Theory.

Keywords

Semiotics, Computational Semiotics, Mental Disorder, Virtual Environments, Philosophy of Mind, Human Information Processing.

1. INTRODUCTION

"Categorizing the sensual modes that bodies can experience fulfills several functions. It elicits a discourse system; it represents efforts to frame the body as an ordered set of impressions which could be disrupted and require re-ordering (implying a structure to do the ordering); and it implies a binary view of the ways that bodily experience is mediated -- the opposition of order and chaos within the frame of a single physicality." [1]

This paper investigates possible interrelations between two symbolic coded concepts of meaning-making, 'madness' and 'virtuality'. The authors know of no further research that explicitly articulates such a liaison between these two social and sensual experiences, both of which seem to be increasingly pervasive in advanced capitalistic societies. By 'pervasive', the authors mean experiences, and workings of these symbolic coded concepts.

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

In the following, we mainly focus on structures, processes and performances of madness and virtuality. Both concepts represent certain semiotic communicative codes, and communication concepts, which we call 'coded acts of practices'.

1.1 Theoretical Propositions

Both madness and virtuality demonstrate scope and elasticity of social and sensual experiences. Thereby, the two coded concepts generate a fluid transgression between form and content, physical and mental. Madness and virtuality pose questions about what is real, and what not, and mostly: what is in-between. To seemingly 'stable' physical and mental societies' structures, both queries can be dangerous. Hence, madness and virtuality themselves appear ordered, and disordered, depending on the perspective: A madman may, e.g., experience the world coherently, whereas an outside therapist may experience the opposite.

Contrary, a visitor to a virtual environment may undergo a social and sensual tour-de-force, whilst the computer system performing the virtual environment considers the locale to be in perfect order: The virtual environment works. If it doesn't work, it must be disordered and incoherent, and would therefore be interrupted, or not established at all, or carried out irregularly.

1.1.1 Symbolic Order & Symbolic Chaos

Let us clarify the terms 'order' and 'chaos' as being applied in this paper: Besides an ontological distinction between form and content, both madness and virtuality seem to contain *symbolic order* - constructed and maintained by e.g. relatively rigid hegemonial definitions, dominant rules and values at-present and *symbolic chaos*. In this paper, symbolic chaos - or symbolic 'disorder' when directly opposed to symbolic order - unfolds recessive, non-linear structures to the eyes of the observer.

Should it be proven true that madness and virtuality really challenge coded concepts of reality, at least they share this clash between strange and 'normal' attractors.

These recessive structures, then, need not to be less complex than structures of symbolic order. 'Complex', when used as a term in this paper, signals a multiply encoded, accumulated meaningmaking system's extrinsic-organization, rather than its selforganization.

1.1.2 Physical and Mental Discourses

Once read, the governing hegemonial symbolic order categorizes these recessive symbolic structures as being incoherent insofar that they might attack, withstand, mirror, or reject dominant structures of coded acts of practices. Thus, symbolic order and symbolic chaos - found in symbolic coded concepts of meaningmaking - constantly negotiate meaning within these meaningmaking systems. Such negotiations are carried out by human and computational performances during physical and mental discourses. The result of these performances are describable coded acts of practice.

If a discourse was a war of differering meanings, then symbolic order, and symbolic chaos, would battle bi-frontally over representations of both physical form and mental content, that is, representations of madness and virtuality.

Additionally, madness and virtuality themselves can be understood as prototypical examples of symbolic coded concepts which are not only physical, but at the same time mental. The two categories transgress into each other.

1.1.3 Qualities of Madness & Virtuality

Madness and virtuality possess qualities of structuring, processing, and performing information. From another perspective, madness and virtuality also mediate information about themselves, apart from mediating information about the world shaping the way these two coded concepts are perceived.

1.1.4 Points of View

Our brains seem to have internal theories about what the world is like not only in terms of mere perception. Here, points of view occur individually, because of different gender, age, ethnicity, race, and habitus, to name a few indicators of social identity.

A similar, inverted interpretative code can also be examined in external, physical points of view *onto* the workings of the mind: A computational semiotician will look at them differently than, say, a philosopher of mind (see Figure 1), a neuropsychiatrist, or a computer. This computer that e.g. generates virtuality in form and content of a virtual environment, and a madwoman share a mutual mode of self-reflexion: From the point of view of the authors, to a certain degree both computer and madwoman are not reflective at all when thinking about their representation of madness, or virtuality.

1.1.5 Research Approach

At this point, we need to draw attention to our research approach towards 'madness' and 'virtuality': The authors analyse these terms according to their conceptual quality of being semiotic representations of meaning-making, not in terms of medical diagnosis, for example.

1.2 Definitions

In this section, we present definitions of concurrent terms used in this paper.

1.2.1 Symbolic Signs

Within our context, symbolic signs designate historiocritical coded concepts built of arbitrary signs [2] which negotiate meaning in physical and mental discourses without having been learned, or are just experienced without outer reference, see e.g. [3].

1.2.2 Signifiers & Signified

Symbolic signs themselves are put together of signifiers, and signified. This structure might be compared to the societal agency-agent-paradigm [4], including two major differences signalling a semiotic approach: historioicity of the structure, and diverse and dynamic points of view: Form and content of a structure change over time, depending on who generates, and who receives form and structure.

1.2.3 Upper-Level Signifiers & Signified

Symbolic signs of an upper hierarchical level consist of an upper level signifier and upper level signified artifacts.

1.2.4 Meaning

Meanings - expressed through physical and mental definitions - of these symbolic signs differ over time, too - linguistically, culturally, and perceptively, and insofar socially and sensually. That is, meanings install a social and sensual definitorial hegemony of a certain symbolic reading.

1.2.5 Meaning-Making

Eventually, clusters of complex symbolic signs form a meaningmaking *system* with an intrinsic symbolic order - imagine a puzzle made out of pieces, that pictures a part of a brain, e.g. the optic tract. The finished jigsaw is just another piece in a bigger puzzle that would eventually show the whole of the brain. However, this brain is just another piece in an even bigger jigsaw. We assert that computers and their practical performance can be de- and reverseconstructed similarly. Meaning-making is characterized by digital and analogue codes mapping reality; a discourse constructs and maintains forms of reality whilst reflecting symbolic order or symbolic chaos inscribed into meaning-making. This differentiates pragmatic meaning-making systems from semantic or syntactic systems, where e.g. semantic research focuses on the preferred reading rather than the representational performances of systems [5], and thence the resulting codes acts of practices.

1.2.6 Performance

In social sciences, 'performance' indicates one *coded act of practice*. The interpretation of symbolic signs is determined by signs' relevance and agreement throughout other participants in a physical or mental discourse. Participants of a mental discourse can, e.g., be pre-existing symbolic coded concepts.

However, it is not clear whether - in a Hegelian sense - there is an ultimate reality (or an ultimate meaning-making system). Meaning-making systems could reside - according to Nietzsche before he fell for eternal recurrence - in a state of flux, and hardly be explained and analysed: They themselves, then, would present no facts, but only interpretations.

1.2.7 Codes & Mappings

Mappings and codes feature analytical semiotic agendas in order to approximate symbolic coded concepts of meaning-making. In this context, mappings operationalize symbolic concepts, and help to approximate these, whereas codes organize symbolic signs into meaning-making systems. We apply these semiotic agendas to concepts of madness, and concepts of virtuality.

1.3 Computational Semiotics and Madness

As outlined in 1.1.4, not only physical and mental symbolic orders of representations can be semiotically prescribed, but also physical and mental (interior) symbolic chaotic perceptions can be matrixed semiotically. Apart from perception, computational semiotics know of other fields of e.g. interior investigation, for example behavior generation.

Thus, in this paper we suggest to examine aspects of a symbolic coded concept of meaning-making - madness - more closely, so as to apply those perceptions of versions of reality upon comprehending virtual environments, and vice versa. In other words: We do believe it is possible to learn from codes and mappings of madness in order to create virtual environments, and learn from codes of virtual environments in order to understand madness.

1.3.1 Extrinsic-Organization of Self-Organization

As shown above, symbolic constituencies do not need to be chaotic inherently. At times, they rather show a certain - though unfortunately pathological - coherence, pattern, which may signal a complex symbolic order, just like a flock of seagulls fluttering into the skies, suddenly changing direction as if following a secret path. This procedural behavior of intra-social systems is neither novel to neuro- and cognitive science, nor to social sciences. The latter have started to study dissipative systems most recently, for example in [6]. Still, those studies bear in mind that they monitor these systems with a certain point of view of their own: The studies extrinsically organize the systems' self-organization.

When browsing research on simulation of reality, and simulation of complexity - the two major streams of intentional and planned virtual *expression* in immersion it seems - researchers face similar problems. Inscribing reality, and especially complexity, into virtual spheres is not a simple task. In the field of volume rendering in immersive Virtual Environments [7] [8], research e.g. mimicks complex processes through literal visual magnification and/or multiple layering of voxels. However, visitors to these worlds are voluntarily interiorised into a system of symbolic negotiation like they were temporary sucked into a black hole.

Other than this, real face-to-face communication channels such as para-verbal or non-verbal ones [9], are only fairly compensated especially in collaborative virtual scenarios. Besides their procedural act, performative practices of these channels follow a system of symbolic order and symbolic chaos. In both cases, seemingly ordered 'real' symbols and seemingly chaotic, or 'virtual' symbols compete with one another, and therefore emerge as seemingly chaotic or seemingly ordered, measurable performances on the surface of reality, or within virtuality.

1.3.2 The Social Dimension of Meaning-Making

The inscription of an extrinsic-organization into the selforganization of both madness and virtuality wards both phenomena sociability, which makes it only possible to analyze the phenomena semiotically. Structures, processes, and performances of these phenomena are socially engineered, whilst these phenomena change dynamically during this act of social engineering.

The two symbolic coded concepts in this paper work as versions of reality, because they challenge, and substitute reality. Beyond this, both coded concepts become more and more pervasive in society. In this paper, we assume that this procedure leads to a transgression of the coded concepts, once they establish themselves in our perception of reality.

1.4 Re-questioning our topic

We are aware that above mentioned assumptions map a much broader ethical and moral area of thought, an area we have not at all covered. For developers, the pain and heartbreak relatives of patients, or patients themselves, have to suffer from, is maybe understandable rationally, but not emotionally. Especially, when those developers have been unaffected by meaning-making systems of madness before.

In this discussion, we will try and concentrate on the possibilities granted by the gift of experiencing the world differently, of being "mad" (and being "virtual", too).

2. REPRESENTATIONS OF MEANING-

MAKING

In the following section, we discuss two symbolic coded concepts of meaning-making, madness and virtuality. We call both 'representations', alluding to their overall social & conceptual appearance as discussed above.

In section 3, we will then explain how come we found evidence of a possible comparison, following which we will present this comparative discussion of semiotic aspects of virtuality and madness, concentrating on "mappings", and "codes". We conclude briefly with future research directions, and suggest a framework of how to apply findings of this paper upon building Virtual Environments.

We want to emphasize that our paper represents a discourse by itself, an approximation of the topic.

2.1 Representations of Madness

'Lunacy', 'insanity', 'dementia', 'mania', 'frenzy', 'mental defect', 'mental disability', 'mental disease', 'mental disorder', 'mental deficiency', 'mental disturbance', 'mental derangement', and 'mental illness' format discursive representations trying to describe and therefore categorize phenomena into symbolic order, held against symbolic chaos. We will cluster these terms with the representational term 'madness' henceforth.

Schools within social anthropology, cultural studies and the philosophy of everyday-practice and meaning-making systems [10] – 'meaning-making' as outlined in our introductory equation - identify phenomena not stringently according to quantitative measures. Rather, schools concerned with meaning-making practices detect the symbols, or constructions, humming underneath definitions qualitatively. Styles of this kind of research and interpretation offer empirical diversity, be they sociohistorical, structuralist [11], descriptive [12], narrative [13] or iconological. As part of a well-established tradition, most anthropological researchers bear in mind that by deconstructing semiotic codes, the discourse model itself opens up new understanding of the codes, and encodes new meaning. Therefore, their personal impact into the research is reflected. We shall return to our personal interest in section 3, "Mythology". By drawing an analogy from cybernetics, let us clarify aforementioned discursive representations. Those set up a cluster model of the world, which

also represents a system within other systems. A meaning-making system has some knowledge if this system carries a code of some part of reality as it is perceived by the system. In this sense, symbolic order and its counterpart symbolic chaos should be understood and taken for complex symbolic systems of symbolic signs on a spectrum, which can be expressed by either language – 'lunacy', 'mental illness' and so forth – regardless of their exact neuro-psychological or cognitive definitions, overlaps and differences as being taught at universities worldwide. Or, those symbolic systems manifest their presence through e.g. software applications' everyday-practices, or development, and last but not least, art.

2.2 Everyday-Practices of Madness

Specifically madness has brought to life outstanding pieces of art. In Vienna in 1997, an exhibition entitled "Kunst und Wahn" -"Art and Madness" - traced this exceptional relationship between expressive imagination and psychosis [14]. [15] refers to historic cycles bearing power relations shaping these relationships.

So on the one hand, in order to approximate 'madness' from an semiotic standpoint, a thorough investigation would need to ask for the power relations wherein madness is brought into our perception. We will not conduct this kind of research in this paper.

On the other hand, whilst generative mechanisms underneath this representation would be worth a look at, interrelationships and between representations of meaning-making are of interest, specifically their performances. Comparisons may be a useful method to come to closer terms with both clusters.

We seize our following attempt in reflection of a *gedanken-experiment* that has been revolving around common conversations about human-computer interfaces in immersive Virtual Environments - we presume CAVEs and CAVE-like installations are known across our readers. If not, visit web sites such as e.g. [16] or read [17], where those are introduced.

In our paper, madness is understood as a symbolic performance embedded into a meaning-making system which is encapsulated and embraced by exogene, clustered definitions and endogene symbolic-chaotic patterns forming systemic subsets. Both poles live on a symbolic spectrum, though their transition into one another - that is, from observation and definition to proprioception - shifts blurry, seamless and reciprocally. One may say that this kind of spectrum could be called a dynamic transgression of madness, where ordered and chaotic symbols of meaning-making resort into one another.

Accordingly, Michel Foucault has been describing such a transcending bipolar archaeology of madness in-depth [18].

In toto, we can state that representations of madness aggregate a symbolic meaning-making system where representations of symbolic order and symbolic chaos negotiate meaning during their performances.

If madness were an alternate version of reality – which it is for the ones suffering from it - it can be compared to virtuality, which by its name declares it is not real, but virtual.

2.3 Representations of Virtuality

Virtuality "in an everyday-practice administers non-Cartesian,

three-dimensional Internet-based, computer-generated coordinate environments" [19], although the authors have been working in immersive, network-based Virtual Environments lately. An updated definition matched to this paper would state that virtuality, in an everyday-practice, administers immersive non-Cartesian, three-dimensional, network-based, computer-generated and –rendered coordinate environments, autarkic or collaboratively, but yet physically inhabited. 'Being there' - which in usability studies concerned with immersive virtuality placeholds modes of presence and awareness (and, we might add, proprioception) - fosters a dynamic spectrum of order transgressing into chaos, and contrariwise.

Disruptions of this outside order - maintained by a visualization system – lead to a lessened intra-perception of 'being there', which can be claimed symbolically chaotic. Similar to the preceding section, a representation of virtuality exists as a symbolic meaning-making system. A comprehensive and more detailed overview of models partaking into this system gives [20], though contrary to our semiotic approach, this study engages with usability engineering of Virtual Environments.

Virtuality as a symbolic meaning-making system produces expressive forms, e.g. simulations of complexity and reality; therefore, virtuality does not simply equal simulation, see [21], especially when referencing to semiotics.

Quite interestingly, long-lasting architectural definitions, such as outlined in [22], try to lift-off real space from "mathematical space" by attributing the latter with transgressing boundaries, homogeneity, infinity, and a free-floating coordinate axis. For our case, the following subsection shows an eventual attempt of defining a representation of virtuality.

2.4 Everyday-Practices of Virtuality

Virtuality, in an everyday-practice, administers immersive non-Cartesian, three-dimensional, network-based, computer-generated and –rendered coordinate semiotic environments, autarkic or collaboratively, but yet physically inhabited. The inhabitation is depending on a visitor's maximum feeling of presence, awareness and her proprioception in this environment which enforces transgressing boundaries, inherent homogeneity, infinity, and a free-floating coordinate axis.

Symbolic signs within this environment can be simulations, or other pragmatic expressive orders. They trigger the degree of perceived inhabitation, and sustain bodily experience through a multitude of signs. Virtuality as a whole aggregates a semiotic meaning-making system where symbolic order and symbolic chaos negotiate meaning, but only if virtuality is in fact inhabited: A semiotic representation of virtuality comes only into existence when a visitor, and the steering computer, form a temporary symbiotic entity that characterizes virtuality.

3. MYTHOLOGY

In the following section, we describe how we found evidence of a possible comparison between madness and virtuality. We use artistic pieces representative of representations of madness, and virtuality, to exemplify interfacing aspects of the two concepts.

Furtheron, we reference to other representations of madness, and virtuality that triggered our interest in a comparison.

3.1 Art & Madness

Whereas [23] renders an visual iconography of psychosis, [24] emphasizes on a genre within pictorial art, 'L'art brut', or Raw Art in the United States. Other languages of art have likewise brought madness into reality, either through their originators, or their products, or both. In literature, evidence of originators' madness can be traced in Friedrich Hölderlin's oeuvre. Works spieling with the meme of madness are numerous, just think "Die Blendung" by Elias Canetti, Georg Büchner's "Lenz", or Sylvia Plath's "The Bell Jar". Most recently, the movie "Girl, Interrupted" dealt with a representation of madness, and even music group's names refer to the theme – think British band 'Madness'.

Musical and applied art are widely spread therapeutic means by which patients are stimulated and helped to express themselves. We encountered a thick red line that indispensably ties imagination and creativity to madness, and encodes an artistic discourse dealing with boundary experiences, and creativity. Even if not considered a "mad" artistic expression, the cited examples picture artistically how madness as a symbolic concept challenges versions of reality.

3.2 Art & Virtuality

The University Stuttgart based CAVE-like Virtual Environment has been utilized for artistic experiments over and over again. Most artistic endeavors challenge perceptions of external reality by triggering a Look & Feel that does not match with what Jean-Pierre Changeux describes as the "anthropic principle" [25]. Instead of anthropocentrically assigning "specifically human qualities to external reality" [26] - which would force Virtual Environments to anthropomorph - visitors of artistic Virtual Environments tend to machinamorph insofar that they adapt to a computer-controlled world.

This discourse seems remarkable, since it underlines and supports the immersive quality of CAVE-like Virtual Environments when confronting human beings with rules and constraints not known to outside meaning-making systems. Here, similar to 'art-on-theedge', virtual symbolic chaos is held against an anthropic symbolic order of reality. On the other hand, one could argue that in order to understand exactly what happens to a machinamorphed visitor within an artistic Virtual Environment, that above mentioned anthropic principle would just keep ready the key to unlock the door to first-person science, where a perceptive methodology "based on a relinquishing of memory and imagination in favour of observation" [27] nullifies any *a priori* meaning-making system. Within this context of a meditative, ineffable and almost autotelic being, semiotics would have come to an end.

3.3 An Architect and an Anthropologist on Mars, dreaming

Besides artistic expression - where madness and virtuality can in fact transgress into one another - we remembered Oliver Sack's popular volumes [28] [29], where he describes some of his patients' representations of madness. Reading through the tragic, astonishing and sometimes pleasing episodes, we encountered something like 'guidelines': Wouldn't it be challenging to transfer those losses, amplifications and journeys into Virtual Environments?

We resembled Alan Lightman's "Einstein's Dreams" [30], too, which had almost the same intention as Sacks – to describe somebody else's would-be experience, and perception of space and time. Lightman's novel renders virtuality as a twosome literary simulation of reality and complexity – what would happen to space, time and everyday practices in alternate physical worlds if they were for real?

Sacks, on the other hand, literarily renders madness as if his patients' cognitions were on the boundary of experience, as if they were virtual: He describes alternate versions of everyday practice (but not at all simulations). Contrary to [31], where social collaboration is held responsible for creativity, Sacks' patients' alternate versions of everyday practices are immanently and individually creative - if the pathological disease precipitating creativity is not taken into account as a, if you want, silent collaborator.

We took this thesis on, and started to ask us to what extent representations of virtuality and madness correlate, since the narrations of both Lightman and Sacks seemed to match at least discursively.

3.4 Computers and Madness

Simplistically put, scientific research in the field of computing machinery either concentrates on regulating madness' surrounding organizational structures, for instance fiscally and administratively [32]. Or, which is of more influence for this paper, constructionist approaches in complex adaptive systems seem handy enough to apply them to our representations of madness and virtuality.

The computer is in control of virtuality, as opposed to a "decentralized mindset" as suggested by Mitchel Resnick [33], where he states that "randomness can help create order", and ongoing: "In many self-organizing systems, random fluctuations act as the seeds from which patterns and structures grow" [34]. The implicit question we raise here is whether symbolic randomness – a superficial sign of potential complexity – equals symbolic chaos as we defined it in the preceding sections.

We will not answer this question, since this section intends to show what route we took, but suggest some questions instead: Can random parameters encourage virtual representations of madness? If yes, can computational semiotics adopting behavior generation - see section 1.2 – induce representations of madness into digital codes and therefore evoke creative behavior?

3.5 Computers and Virtuality

Daniel Weiskopf of the University Tübingen has been working on a project called "Virtual relativity" [35], which simulates the complexity of special relavitiy in a CAVE-like immersive Virtual Environment. Weiskopf has, analogically spoken, visualized some chapters of Lightman's novel. He has also set forth a discourse where an anthropocentric view of the real world is undermined by a simulation of complexity.

From a rationalist's view, knowledge of special relativity will explain being within this meaning-making system fully. From a constructivist's angle, a visually chaotic structure is put into symbolic order. From an architectural standpoint, this mathematical space is non-existent in reality. From a anthropological standpoint, Virtual Relativity performs a boundary experience in which we participate for real – the application sets up a symbolic chaotic meaning-making system, one must almost dare to say: a mad meaning-making system.

From the steering computer's perspective, this all doesn't matter. If the machine which renders special relativity were a brain, it wouldn't be aware of what it is doing to the extent that it doesn't gain positive feedback following its actions. It just *is*, and therefore has relinquished a priori memory and imagination in favour of observation ('a priori' implying the state before Virtual Relativity is launched). Average visitors to this virtual world transgress from symbolic order to symbolic chaos, but form in affiliation with the 'brain' a temporary, symbiotic-like entity: The control-less attendants now possibly face a representation of madness.

We considered this supposition the most thrilling aspect patronizing a comparison between representations of madness, and virtuality.

4. MAPPING CHAOS, CODING ORDER – CODING CHAOS, MAPPING ORDER

This section features two conferring semiotic aspects of representations of virtuality, and madness, mappings and codes.

4.1 Semiotic Mappings

Methodological mapping stems from communication studies, and obviously, from geographical sciences. In semiotics, mapping has grown into a research style which helps to operationalize and approximate a topic. It can be a survey's starting point for an analysis. We had the fortune to find real maps concerned with both madness and virtuality, which seemed promising, resulting in the following mapping of the topics, exemplified by maps.

The art of map-making has entered both representations of madness and non-immersive representations of virtuality a long time ago already. Whereas organic matter is, in the former case, signified by the signifying geographical localization of neurological diseases - e.g. neuroimaging [36] - anorganic relationships present the latter's signified, with signifiers in form of visualizing spatial notions of these relationships, see [37].

4.2 Maps

The symbolic sign, which already pictures a meaning-making system, is in both cases a map. A multitude of maps (signified) showing different renderings and versions of spatial relations and processes (signifier) establish an upper level sign: an atlas. This atlas, then, has some knowledge if this atlas carries a model of some part of reality as it is perceived by the whole atlas. The 'text', as some semioticians even term a map, mediates information about locale, function, communicative acts, geometrical frames, distance, directions, topology, routes etc. of both representations. These atlases facilitate an understanding of madness and virtuality spatially – they establish symbolic order.

Atlases of virtuality also enforce the prophecy of virtual immersion: They engage their users to believe in the cultural, but yet technical spatiality of non-immersive representations of virtuality, whereas maps of madness allude to the naturalness of healthy, ordered brains. Both maps intend to put symbolic chaos into symbolic order, but require special skills to be decoded.

If maps signifying representations of madness, and those of virtuality, are compared, it intrigues how much they relate in terms of cognitive map knowledge. But the most important factor - a converging sign for both kinds of maps - stays navigation. Madness, likewise virtuality, can be navigated, and steers navigation. According to our definition in subsection 2.4, representations of virtuality in our case take place in immersive Virtual Environments - so what does a map of immersive virtuality look like?

We suggest that the Virtual Environment *itself* configures a map: If a map delivers complex and abstract information concerning navigation in a certain environment, then it must be true that immersive Virtual Environments represent maps of virtuality to visitors – the latter become part of the map-landscape. On the contrary, maps of madness do allow for a exterior surface navigation of madness merely, but not for a first-person perception of the meaning-making system at work inside the patients' brains. Performances - or coded acts of practice – appear to be navigational practices in both representations, then.

4.3 Semiotic Codes

Codes organize symbolic signs into meaning-making systems. When codes are used in a certain familiar chronological sequence, they can be interpreted by both originators and recipients, with reference to their appropriateness – semiotic codes are the wind in the wind chime.

Virtuality uses digital codes to convey information; representations of madness use neural analogue codes, yet the transmittal processes are bio-chemically maintained. What is important to us here is the mutual inherent *structure of articulation* within both codes.

4.4 Coded Articulation

Virtuality can be broken down into tiniest pieces of articulation easily: Bits and bytes of the programming code, which do not possess meaning by themselves. Being a minimum functional unit, a bit lacks meaning, though one may argue that a bit surely has to decide whether it wants to be zero, or one, and therefore carries meaning, because it equals either true, or false. If these two categories are numeric states, they can only be influenced from the outside, not by the bit itself. Therefore, a bit does not carry meaning without exterior manipulation. Though, if a bit could be both zero and one at a time - which it can - semiotics would probably feel a bit better about this paradigm.

Only in combination with other referencing bits, the sole bit gains meaningfulness. These combined bits become a symbolic sign, which, signified by other byted bits, form a piece of software running inside a Virtual Environment. Here, articulations of the digital code determine a direct symbolic order, whereas the expressive performance of these chunks of code can be of the complete opposite. Ordered software pieced together out of single bits may visualize complexity, or chaos. Imagine a waiter who serves your favorite pasta dish, which looks and smells mouthwatering, and has been cooked exactly to your order. Before you notice, the waiter flips the warmed porcelain plate a 180° - there you go!

Such an anachronism can also be detected in representations of madness. Although most representations of madness are pathologically dynamic, we assume they still contain a basic anthropological element, perceptual constancy. Citing [38], perceptual constancy as an analogue code ensures that "the

variability of the everyday world becomes translated by reference to less variable codes". Most of Sacks patients held such a perceptual constancy, but one that can be described as an *alternate representation* of at least the authors' common perceptual constancy. The anachronous element with representations of madness manifests itself in the complete reverse semiotic transgression from symbolic order to symbolic chaos. Now, your dish consists of a huge rubber tire, but a candle-lit dinner table awaited you, and the waiter dances towards you in rhythm with the gentle melodies of a lounge combo. Everything, apart from the food you eat, seems perfect, and comfortable.

The structure of articulation in representations of madness doesn't ask for appropriate codes; it sets up its own, consisting of phrasings, variations, shifts, or inventions of codes. "Randomness can help create order", Resnick writes, as can slightly "bracketed" chunks of code, whilst building "bracketed perception" piece by piece [39].

In comparison, codes within virtuality need a visitor to generate meaning. Codes within representations of madness make up meaning by themselves. Yet, if Virtual Environment's visitors form temporary entities with the computer, they do experience almost the same as some of Sacks' patients: They are not in control of navigating through a mapped landscape. Since people in CAVE-like environments differ in their interpretation of symbolic signs, and the codes brought along with these, they will experience differing modes of awareness, presence, and proprioception, ditto depending on the kind of Virtual Environment.

As opposed to Sacks patients, regular visitors to Virtual Environments will have a feeling of self-reference because of code practices they brought with them from the outside; in o

ther words: they are able to *compare* different meaning-making systems, like they would compare their wristband watch's time with the present time zone after a long flight.

5. CONCLUSIONS

In the beginning of our paper, we stated that we did believe it would be possible to learn from codes and mappings of madness, and virtuality, and apply findings to each of the pair. Virtual Environments, on the one hand, could help compensating patients' *losses*, or weaken their experienced *amplifications* [40].

On the other hand, let us present a buzzword list of our key findings in order to support our argument:

- Representations of madness, and representations of virtuality seem to show general conformities in the way they are practiced, and practicable
- Performances or coded acts of practice appear to be navigational practices in both representations
- Codes within representations of virtuality need a visitor to generate meaning; codes within representations of madness make up autarkic meaning
- Both themes bear transgressional character, stepping from symbolic chaos to symbolic order, and vice versa
- Representations of madness seem to present alternate, "bracketed" versions of common perceptual constancy, likewise representations of virtuality

- Both madness and virtuality work as concepts where physical and mental interconnect: The physical becomes mental, and the mental becomes physical
- Through virtuality, representations of madness find their way into reality; through madness, representations of virtuality as we may build them, are already lived
- Both concepts challenge symbolic orders of reality

6. FUTURE RESEARCH

We have shown that madness and virtuality have a lot in common in the ways they challenge coded concepts of reality. To what extent they can inform each other purposefully (and possibly therapeutically), would need to be tested in future Virtual Environments, really – in co-operation with and between patients, therapists, and researchers.

The social dimension of interrelations between madness and virtuality, and their impacts onto physical and mental societies will need to be explored much more deeply, too.

7. ACKNOWLEDGEMENTS

This paper is partially funded by the Universitätsbund Tübingen, and the HLRS at the University of Stuttgart. The authors wish to thank Caroline Sexauer, Frank Dehner, Oliver Schippin, and Jürgen Schulze-Döbold, Uwe Wössner and Dr. Ulrich Lang (High Performance Computing Center, University of Stuttgart, Germany).

8. REFERENCES

- Stone, Allucquere Rosanne (1993), VIOLATION AND VIRTUALITY: Two cases of physical and psychological boundary transgression and their implications, see http://www.actlab.utexas.edu/~sandy/violation-andvirtuality.
- [2] Barthes, Roland (1972), *Mythologies*, Hill & Wang, New York.
- [3] Peirce, Charles Sanders (1958), Collected Writings. Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks, Harvard University Press, Cambridge, MA.
- [4] Minsky, Marvin (1988), *The Society of Mind*, Touchstone, New York et al.
- [5] Morris, Charles W. (1970), Foundations of the Theory of Signs, Chicago University Press, Chicago.
- [6] Harvey, David L. and Michael Reed (1997), "Social Science as the Study of Complex Systems", in *Kiel, L. Douglas and Euel Elliott (eds.), Chaos Theory in the Social Sciences. Foundations and Applications*, The University of Michigan Press, Ann Arbor, pp. 295 – 323.
- [7] Hibbard, Bill (2000), "Visualization Spaces", ACM SIGGRAPH Computer Graphics, Vol. 34, Nr. 4 (=November 2000), pp. 8 – 10.
- [8] Schulze-Döbold, J., U. Wössner, S. P. Walz, U. Lang (2001), "Volume Rendering in Virtual Environments", Proceedings of the Joint Fifth Immersive Projection Technology Workshop and Seventh Eurographics Workshop on Virtual Environments, Springer, Vienna et al., In Print.

- [9] Goffman, Erving (1959), *The Presentation of Self in Everyday Life*, Anchor Books Doubleday, New York et al.
- [10] Bourdieu, Pierre (1977) *Outline of a Theory of Practice*, Cambridge University Press, Cambridge.
- [11] Foucault, Michel (1988), Madness and Civilization. A History of Insanity in the Age of Reason, Vintage Books / Random House, New York.
- [12] Lorenz, Thorsten (1990), "Die Psyche zählt statt erzählt. Zur Entdeckung einer kinematographischen Mathematik der Seele", in *Hörisch, Jochen and Michael Wetzel (eds.)*, *Armaturen der Sinne*, Wilhelm Fink Verlag, Munich, pp. 247 – 264.
- [13] Stone, Allucquere Rosanne (1993), see [1]
- [14] Brugger, Ingrid, Peter Gorsen and Klaus Albrecht Schröder (eds.) (1997), Kunst und Wahn. Ausstellungskatalog. Kunstforum Wien, Dumont, Cologne.
- [15] Vojvodík, Josef (1998), "Iconography of Psychosis", Atélier Nr. 4, February 19, 1998, Lecture Series at The Center for Contemporary Arts, Online Synopsis available at http://www.fcca.cz/ENG/calendar/atel4.html
- [16] Benokraitis, Andrius et al., "VT CAVE Information Center", Online Document available at http://www.cave.vt.edu/
- [17] Lang, Ulrich et al., "Virtual Environments Lab", Online Document available at http://www.hlrs.de/organization/vis/velab/
- [18] Foucault, Michel (1988), see [11].
- [19] Walz, Steffen P. (2000), "Symbiotic Interface Contingency: The reciprocal emergence of use and abuse", *Proceedings of the Second International Conference on Cultural Attitudes towards Technology and Communication*, Murdoch University Press, Perth, pp. 125 – 144.
- [20] Gabbard, Joseph L. and Deborah Hix, A Taxonomy of Usability Characteristics in Virtual Environments. Deliverable to Office of Naval Research. Grant No. N00014-96-1-0385, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Online document available at http://csgrad.cs.vt.edu/~jgabbard/ve/taxonomy/
- [21] Esposito, Elena (1998), "Fiktion und Virtualität" in Krämer, Sybille (ed.), Medien, Computer, Realität. Wirklichkeitsvorstellungen und Neue Medien, Suhrkamp Frankfurt am Main, pp. 269 – 296.
- [22] Bollnow, Otto Friedrich (2000), *Mensch und Raum*, 9th Edition, Kohlhammer, Stuttgart et al.
- [23] Vojvodík, Josef (1998), see [15].
- [24] Freeman, Barbara (1998), "Alfonso Ossorio, Jean Dubuffet and Art Brut in America", *Raw Vision #17*, Article excerpt

available online at

http://www.rawvision.com/back/ossorio/ossorio.html

- [25] Changeux, Jean-Pierre and Alain Connes (1995), "The Real and the Rational", p. 200, in *Conversations on Mind, Matter,* and Mathematics. Edited and translated by M. B. DeBevoise, Princeton University Press, Princeton, pp. 179 - 209.
- [26] ibid. p. 201.
- [27] King, Mike (1999), "The New Metaphysics and the Deep Structure of Creativity and Cognition", *Proceedings of the third Creativity & Cognition Conference*, ACM Press, New York, pp. 93 – 100.
- [28] Sacks, Oliver (1987), *The Man Who Mistook His Wife For a Hat*, Perennial Library / Harper & Row, New York.
- [29] Sacks, Oliver (1995), An Anthropologist on Mars, Picador, Basingstoke and Oxford.
- [30] Lightman, Alan (1994), Einstein's Dreams, Warner Books, New York.
- [31] Fischer, Gerhard (1999), "Symmetry of Ignorance, Social Creativity, and Meta-Design", *Proceedings of the third Creativity & Cognition Conference*, ACM Press, New York, pp. 116-123.
- [32] Gorodezky, M. and J. Rusnak (1990), "Increasing the Clinical Relevance of a Mental Health Information System", *Proceedings of the conference on Computers and the quality* of life, ACM Press, New York, pp. 93 – 96.
- [33] Resnick, Mitchel (1994), Turtles, Termites, and Traffic Jams. Explorations in Massively Parallel Microworlds, MIT Press, Cambridge, MA, and London.
- [34] ibid. p. 137.
- [35] Weiskopf, Daniel (2000), Virtual Relativity, Online document and downloadable beta version of "Virtual Relativity" available at http://wwwvis.informatik-unistuttgart.de/~weiskopf/relativity/vr.html
- [36] Johnson, Keith A. and J. Alex Becker (1999), *The Whole Brain Atlas*, Online document available at http://www.med.harvard.edu/AANLIB/home.html
- [37] Dodge, Martin and Rob Kitchin (2001), *Mapping Cyberspace*, Routledge, London and New York.
- [38] Nichols, Bill (1981), Ideology and the Image: Social Representations in the Cinema and Other Media, Indiana University Press, Bloomington, p. 26.
- [39] Resnick, Mitchel (1994), see [34].
- [40] Sacks, Oliver (1987), see [28].

A Missing Link: The Role of Semiotics in Multiagent Enviroments

Mirko Petric Department of Visual Communication Design UMAS, University of Split Glagoljska bb 21000 Split, Croatia

mirko.petric@umas.hr

Inga Tomic-Koludrovic Department of Sociology Faculty of Philosophy, Zadar UMAS, University of Split Glagoljaska bb 21000 Split, Croatia

inga.tomic-koludrovic@umas.hr

Ivica Mitrovic Department of Computer Science, FESB UMAS, University of Split Glagoljaska bb 21000 Split, Croatia

ivica.mitrovic@umas.hr

ABSTRACT

Approaches to artificial intelligence research based on autonomous agents have thus far primarily relied on cognitive science and other psychologically motivated approaches to the social world. These approaches have been a logical outcome of the researchers' concentration on agent building. However, recent developments in the field have revealed the need for a wider and primarily socially centered interpretive framework in which to account for or model the agents' behavior. Dautenhahn's attempt to interpret the agents' social behavior has been based (frequently only implicitly) on macro-sociological approaches and biologically motivated evolutionary determinist theories. In order to make autonomous agents truly "socially intelligent", researchers will sooner or later need to take microsociological theories into consideration. The fact that these theories are interaction and communication-centered opens up the possibility for a relevant application of semiotic theories in the field of AI.

This paper outlines the elements of several classical semiotic theories potentially useful in current and future AI research centered on autonomous agent modeling. Briefly discussed are Greimas-Courtes's distinction between actor and actant, Tartu-Moscow school notions of semiosphere and culture as secondary modeling system, Jakobson's notion of the functions of language, and Eco's semiotic approach to semantics, as expressed by his notion of cultural encyclopedia.

General Terms

Theory.

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands.

Keywords

Artificial intelligence, behavior-based AI, autonomous agents, socially intelligent agents, multiagent environment, semiotics, microsociology, actor, actant, semiosphere, secondary modeling systems, Jakobson's communication model, cultural encyclopedia.

1. INTRODUCTION

Recent developments in the field of AI have brought to the fore the need to account for the social aspects implicit in user-agent interaction or agent interactions taking place in artificial environments without human control. Indicative of the "social turn" in the field is the use of terms such as "socially intelligent agents" and "social embeddednes", as well as the titles of journals, research projects, and symposia¹. Although some researchers have used social conceptions to describe the interplay of artifical agents in multiagent environments, these descriptions have sometimes been discarded by sociologists as "folks sociology"[17] or thought of as "blindly perform[ing] a one-toone mapping of sociological concepts to computer models"[19]. The task of disciplinary collaboration between the disciplines of sociology and distributed AI has been described as a "painful" one[19].

Approaches to artificial intelligence research preceding its current social focus relied primarily on cognitive science and other psychologically motivated approaches to the social world, which

¹ A major source of information in this research area has recently been *Journal of Artificial Societies and Social Simulation*. In 2001, this journal organized a symposium and published a special issue on the topic "Starting from Society: the Application of Social Analogies to Computational Systems" (vol. 4, no. 1). Other journals and conferences have also presented articles and papers on "socially situated AI", "social intelligence" and "socially intelligent agent-building". A major research project funded by the German Research Foundation ("Socionics") has been set up in order to "start a serious evaluation of sociological conceptions and theories for computer systems", with the aim of "developing intelligent computer technologies by picking up paradigms of our social world"[19].

was a logical outcome of the researchers' concentration on agent building.

Recent attempts of a prominent researcher trying to interpret the agents' social behavior[6] have been based (frequently only implicitly) on macrosociological approaches and biologically motivated evolutionary determinist theories. What is obviously lacking in the field is a sustained effort to interpret user-agent and artificial agent interactions in microsociological terms².

However, in order to make autonomous agents truly "socially intelligent", researchers will sooner or later need to take microsociological theories into consideration. The fact that these theories are interaction and communication-centered also opens up the possibility for a relevant application of semiotic theories in the field.

This paper outlines the elements of several classical semiotic theories potentially useful in current and future AI research and agent modeling. Briefly discussed are Greimas-Courtes's distinction between actor and actant, Tartu-Moscow school notions of semiosphere and culture as secondary modeling system, Jakobson's notion of the functions of language, and Eco's semiotic approach to semantics, as expressed by his notion of cultural encyclopedia. Examples of interaction that illustrate the potential usefulness of theories discussed are taken from the field of behavior-based AI³, and specifically from the Carnegie Mellon University's Virtual Theater Project (the Servant/Master Scenarios)⁴.

2. "ACTANT" AND "ACTOR" AS QUALIFIERS OF SOCIAL AGENCY

In any attempt to interpret the interactions taking place in multiagent environments from a sociological point of view, it is important to use the term "social agency" in addition to "social behavior", which almost immediately implies a more psychological approach and behaviorist stimulus-response model⁵. In order to further qualify this term and gain a more

⁴ Detailed accounts of these interactions as well as general information on the emergence and architecture of believable agents developed within the Oz project can be found in [2], [3], [4], [16] and [18]. These articles and a general presentation of the project are presented at the project web site (http://www.cs.cmu.edu/project/oz/web/papers.html). Details of Master/Servant Scenarios, developed within The Virtual Theater Project, are described in [12] and presented at http://www.ksl.stanford.edu/projects/cait/demos/status.html

⁵ One obvious exception to the psychological-behaviorist coloring of the term "social behavior" is its use in Goffman's micro-

complete understanding of the nature and consequences of interactions taking place in multiagent environments, a semiotic distinction between the terms "actant" and "actor", as explained by Greimas and Courtés[11], can also be introduced into discussion.

Based on Greimas's work in narratology, this distinction primarily relates to an interpretation of semiotic phenomena in the medium of natural language, and not to an interpretation of society. However, social relations are clearly implied in the subject-matter and analytical procedures of narratology and, because of this, the distinction is obviously welcome in microsociological accounts of interactions as well.

Greimas and Courtés state that the term "actor" has gradually replaced "character" and "dramatis persona", extending its use outside the purely literary. This is especially interesting in the field of "behavior-based AI" in which theater terminology ("virtual drama", "interactive drama", "character", "story", "scene", "drama manager") is frequently used. In this context, the distinction between the terms "character" and "actor" can be seen as a shortcut to analytically highlighting the social component implicit in the networks of relations forming in "fictional" multiagent environments.

An even more complete understanding of the social implications of actions performed by "characters" in "interactive dramas" can be gained by introducing a further distinction between the terms "actant" and "actor".

Quoting Luciene Tesnière, from whom they borrow the term, Greimas and Courtés state initially that "actants are beings or things that participate in processes in any form whatsoever, be it only a walk-on part and in the most passive way"[11]. In simple terms, "an actant can be thought of as that which accomplishes or undergoes an act, independently of all other determinations"[11]. From a sociological perspective, the term "actant" is stimulating because it pairs off with the notions of role and status ("actantial role", "actantial status"). As narration (in the sociological context, "social process") unfolds, an actant can take on a number of actantial roles.

According to Greimas and Courtés, actantial status is "that which defines the actant at a given moment of the narrative trajectory, taking into account the totality of its previous trajectory (manifested or simply presupposed)"[11]. Actantial role, on the other hand, is defined as "simply the surplus which is added, at a particular point of trajectory, to what already constitutes the actant within the syntagmatic progression of the discourse"[11].

The distinction translates easily into the artificial behavior-based multiagent environment. Behaviorally animated autonomous agents Gregor and Otto in the Master/Servant interaction can be thought of as "characters" with predefined actantial status and actantial roles implied by that status (i.e. expected at the particular point of trajectory at which they find themselves at the moment when interaction takes place). However, what happens during the

² The need to take into account the microsociological focus in the field has been announced by works such as Agre's 1988 Ph.D. thesis on *The Dynamic Structure of Everyday Life*[1] and by the individualist micro-approach of Edmonds's articles [9] and [10].

³ A detailed account of terms used in the field of "new" or "alternative" AI can be found in [18] and [21].

sociological theory. The term "social agency" has been used as a hallmark of sociological approach since Weber introduced it in his 1921 study *Wirtshaft und Gesellschaft: Grundriß der verstehenden Soziologie.*

interaction is a reversal of actantial roles and a corresponding reversal of actantial status, which turns Gregor and Otto into individualized interaction entities better described by the term "actor".

In contrast with "actant", "actor" is "not only the point of investment of [actantial and thematic] roles, but also of their transformations, since discourse consists essentially of the interplay of successive acquisitions and loss of values"[11].

This definition makes it possible to introduce microsociological approaches into the discussion of interactions taking place in multiagent environments. In the course of an interaction the actor not only demonstrates actantial status and role, but is also "incorporated into the discourse" and "maintained througout the discourse - or at least throughout a discursive sequence - in keeping with the principle of identity"[11]. Through various narrative and semantic investments, the actor acquires the appearance of an autonomous figure in the semiotic universe.

From a sociological point of view, this notion would correspond to a "social actor" manifesting itself as an autonomous figure in "social interaction" (i.e. in society pressuposed by it), and would obviously include Gregor and Otto of the "Servant/Master" interaction.

3. BIOSPHERE, SEMIOSPHERE, SECONDARY MODELING SYSTEMS

Given the current foci of interest in both disciplines, the most pertinent application of semiotic methodology developed within the Tartu-Moscow school to the field of AI would at first glance seem to relate to the current reelaborations of the late Lotman's distinction between biosphere and semiosphere.

Lotman derived the notion of semiosphere from V. Vernadski's biosphere and conceived of it as an abstract space in which languages, texts and cultures intertwine[23]. According to Lotman, semiosphere is the world as defined by the semiotic competence of living organisms. The process of sign generation (semiosis) would be unthinkable without it. However, this does not mean that this abstract sphere is cut off from what does not belong to it. What crystalized as the most important notion connected with semiosphere is the notion of the boundary, through which it is in contact with the non-semiotic and alien semiotic space. Semiosphere is highly porous (i.e. perforated by multiple inner boundaries) which make possible translation of external messages into its internal space and vice versa.

Researchers who approach the notion of semiosphere from a biosemiotic point of view have also come up with the distinction between "vertical semiotic system", which relates to genetic communication down through the generations, and "horizontal semiotic system", which relates to the communication throughout the ecological space[14]. It is through this kind of horizontal communication that semiosphere, conceived of as an autonomous sphere of communication is also a precondition for advanced social complexity and learning processes that bring it about.

Given the embodied nature of autonomous agents and the emergence of social aspects in their interactions, as well as the importance of learning processes in the field, the emphasis put on the social effects of horizontal semiotic communication would seem to possess an obvious significance in multiagent environments. From a sociological point of view, when higher levels of complexity are reached in this environment, parallels could even be drawn between the notion of boundary, central to the notion of semiosphere, and the notion of the boundary as described by Luhmann.⁶

However, the current level of development of autonomous agents imposes restrictions on the use of notions such as biosphere and semiosphere, as well as on the aplication of sociological theories of a corresponding level of generality. Regarding the technical limitations in the construction of agents, there are on the one hand purely textual agents (such as the cat Lyotard, developed within the Carnegie Mellon University's Oz project) with minimal metaphorical "biosphere" and, on the other hand, agents (such as Woggles, also developed within the Oz project) with physical attributes but very limited "natural language" capabilities. In addition to this, interactions taking place in multiagent environments have thus far resulted only in low levels of social complexity.

At the current level of development, an earlier distinction made by the members of the Tartu-Moscow school would seem to be more fruitful as a means of describing and modeling interactions in multiagent environments. This is the distinction between the "primary modeling system" and "secondary modeling systems". In this context, natural language is seen as the primary modeling system in relation to reality, while the "languages" of culture (only some of which use the natural language as "raw material" but all of which can be interpreted by it) are referred to as secondary modeling systems. These systems comprise "not only all the arts (literature, cinema, theater, painting, music, etc.), the various social activities and behavior patterns prevalent in the given community (including gesture, dress, manners, ritual, etc.), but also the established methods by which the community preserves its memory and its sense of identity (myths, history, legal system, religious beliefs, etc.)"[20].

The only area of secondary modeling systems currently applicable to multiagent environments is obviously the area of "various social activities and behavior patterns prevalent in the given community". The behavior patterns with which agents enter interactions to some degree always resemble those involved in human interaction in a particular culture (in the same way as believable agents are always anthropomorphically modeled).

⁶ A footnote is in order to account for mentioning of Luhmann's theory in the context of a paper insisting on the need for researchers to interpret the agents' social aspect from a microsociological perspective and with the help of semiotic theory. Luhmann's theory approaches social systems at roughly the same level of generality as the notions biosphere and semiosphere do. Its general approach is macrosociological but it effectively combines elements of macro- and micro- approaches in its analytical procedures. In discussing the emergence of sociality in multiagent environments, we are currently preoccupied with an analysis of interactions that bring it about and that can in our opinion be best accounted for from a microsociological perspective. When higher levels of complexity are reached, the application of top-down approaches could also be stimulating, especially when they are communication-based as is the case with Luhmann's theory.

Likewise, the emergent social aspects of multiagent environment interactions can be interpreted with the help of methodology developed to account for interactions that take place in exclusively human environments. Given the current level of development of techniques of agent interaction, Goffman's "dramaturgical analysis" would seem to be the most pertinent approach in this regard, because it emphasises and is capable of effectively interpreting non-verbal elements socially meaningful in interactions.

4. FUNCTIONS OF LANGUAGE IN MULTIAGENT ENVIRONMENTS

Roman Jakobson's classical account of language functions[15] was developed within a communication model dealing only with verbal communication, but could be said to be general enough to retain its explanatory value in environments which are neither purely verbal nor exclusively human.

The six function schema is applicable to interactions in which there are at least an addresser (sender) and an addressee (receiver), a message to be shared, a context of the sharing, a channel through which the message is shared and a code which makes sharing possible. Each of these elements determines a particular linguistic function. The emotive (or expressive) function relates to the addresser and conveys his/her attitude towards the message, the conative function relates to the addressee (i.e. to what the addresser wants the addressee to notice or do), and referential function to the context (i.e. to what is being spoken of). The phatic function "checks whether the channel works"[15] and is typical of the use of language which helps maintain a social relation (contact) through ritualized formulas, while the metalinguistic function is used by the addresser and addressee "to check up whether they use the same code". Finally, the poetic function "focus[es] on the message for its own sake" and in this way "promot[es] the palpability of signs"[15], by means of the repetitions of sound values, stresses, accents, as well as associations and ambiguous boundaries between words and phrases. "Self-focused message[s]" [15] are typical of poetic discourse, but also have a role to play in everyday communication.

In the interactions of anthropomorphic animated puppets (Woggles), as described by Loyall and Bates[16], note can be taken of functions corresponding to Jakobson's referential, conative, emotive and phatic function. Woggles refer to their context, address one another, express emotions and communicate in order to maintain an elementary social ritual.

The interactions of Woggles obviously fall short of the complexity of human communication, but the four functions of language evident in them suffice to situate the agents into the field of sociality (which according to microsociological theories emerges as the consequence of interactions).

5. CULTURAL ENCYCLOPEDIAS IN AGENT INTERACTION

One of the premises underlying symbolic interactionist theories postulates a minimum of the common "definition of the situation"[22] in order for an interaction to take place. Translated into the theory of information-based terminology, this would imply the necessity that the sender and the receiver share the same code in order for the message to be transmitted. However, a somewhat more complex definition is needed to describe interactions which include attitudes and values, i.e. social competences.

The necessary elements of a definition potentially useful in interpreting and modeling the social aspects of the interactions of autonomous agents can be found in Umberto Eco's semiotic approach to semantics. In his theory of semiotics[7] and a subsequent account of interpretive cooperation in narrative texts[8], he developed the notion of "cultural encylopedia", by which he describes "a model of socialized competences at a particular point in history, which the dictionary (a model of ideal competences of an ideal speaker) cannot explain in full measure"[5] [translation M.P.].

Interactions taking place in multiagent environments at the current level of development obviously cannot be expected to parallel the complexities of human interaction. In addition to this, the "approximate and sometimes even illogical" constitution of the cultural encylopedia represents one of the major problems in the field of AI[24] [translation M.P.].

In spite of this, it is evident that "emotional engines" of the afore mentioned believable agents have already been modeled in accordance with principles that correspond to the implications of Eco's notion of cultural encylopedia. For example, in the emotional engine of the purely textual believable agent developed within the Oz project (cat Lyotard), "human" is defined as somebody whom the cat dislikes. In the course of interaction, Lyotard gets food from the human and is pet by the human, and the human becomes somebody liked by the cat.[4]. What has changed in this brief interaction is Lyotard's rudimentary cultural encyclopedia of the "human". The fact that the notion of cultural encyclopedia includes the possibility of change is very important in the field of behavior-based AI, because of the importance of the learning process.

6. CONCLUSION

The relevance of the elements of classical semiotic theories outlined in this paper to the field of behavior-based AI is potentially twofold. On the one hand, it is conceptual (they help us understand the nature of interactions taking place in multiagent environments), on the other hand, they could also have practical consequences in the modeling of agents and interactions.

⁷ Attitudes and values are important elements of the approach to semantics based on cultural encyclopedias (in contrast with the previous, more strictly defined "dictionary"-based approaches). It should be said that, at the current level of development, Lyotard's cultural encyclopedia of the "human" in effect consist only of the attitude toward the human. Neverthless, what changes in the human-cat interaction described in [4], can be viewed as a change of cultural encyclopedia and not merely a change of an attitude. Other attributes (semantic markers) can be added to the notion "human" at a later stage of development. Gregor's and Otto's more complex cultural encyclopedias of the terms "master" and "servant" did not change in the course of the "Master/Servant" interaction described by Hayes-Roth, Van Gent, and Huber [12], although the agents reversed their roles and status.

These theories could prove to be an important link between the researcher's concentration on agent-building and the need to account for the emergent social aspects implicit in agent interaction. What needs to be done in order to make them fully effective in this regard is to explore their interrelations with selected microsociological theories and the possibilities of their combined application in the field of autonomous agents.

7. REFERENCES

- P. E. Agre. *The Dynamic Structure of Everyday Life*. Ph.D. thesis. Massachusetts Institute of Technology Artificial Intelligence Laboratory, Cambridge, M.A., 1988.
- [2] J. Bates. Virtual Reality, Art, and Entertainment. Presence: The Journal of Teleoperators and Virtual Environments, 1 (1):133-138, 1992.
- [3] J. Bates. The Role of Emotion in Believable Agents. Technical Report CMU-CS-94-136. School of Computer Science, Carnegie Mellon University, Pittsburgh, PA., 1994.
- [4] J. Bates, A. B. Loyall, W. S. Reilly. An Architecture for Action, Emotion, and Social Behavior. Technical Report CMU-CS-92-144. School of Computer Science, Carnegie Mellon University, Pittsburgh, PA., 1992.
- [5] O. Calabrese. *Il linguaggio dell'arte*. Bompiani, Milano, 1985.
- [6] K. Dautenhahn. The Art of Designing Socially Intelligent Agents: Science, Fiction and the Human in the Loop. *Applied Artificial Intelligence Journal*, 12: 573-617, 1998.
- [7] U. Eco. Trattato di semiotica generale. Bompiani, Milano, 1975.
- [8] U. Eco. Lector in fabula. Bompiani, Milano, 1979.
- [9] B. Edmonds. Modelling Socially Intelligent Agents, *Applied Artificial Intelligence*, 12 (7), 1998.
- [10] B. Edmonds. Capturing Social Embeddedness: A Constructivist Approach. CMP Report 98-34. MMU, Manchester, 1998.
- [11] A. J. Greimas and J. Courtés. Semiotics and Language: An Analytical Dictionary. Indiana University Press, Bloomington, Ind., 1982. [orig. Sémiotique. Dictionnaire raissonné de la théorie du langage. Hachette, Paris, 1979.]

- [12] B. Hayes-Roth, R. Van Gent, D. Huber. Acting in Character. *Proceedings of the AAAI Workshop on AI* and Entertainment, 1996.
- [13] J. Hoffmeyer. The Global Semiosphere. In I. Rauch and G. F. Carr (eds.), Semiotics Around the World. Proceedings of the Fifth Congress of the International Association for Semiotic Studies, pp. 933-936. Mouton de Gruyter, Berlin and New York, 1997.
- [14] J. Hoffmeyer and C. Emmeche. Code-Duality and the Semiotics of Nature. In M. Anderson and F. Merrel (eds.), *On Semiotic Modelling*, pp. 117-166. Mouton de Gruyter, Berlin and New York, 1991.
- [15] R. Jakobson. Closing Statements: Linguistics and Poetics. In T. A. Sebeok (ed.), *Style in Language*, pp. 350-377. Wiley and MIT Press, New York and Cambridge, Mass., [1960].
- [16] A. B. Loyall and J. Bates. Personality-Rich Believable Agents That Use Language. Proceedings of the First International Conference on Autonomous Agents. Marina del Rey, California, 1997.
- [17] T. Malsch. Die Provokation der 'Artificial Societies': Warum die Soziologie sich mit den Sozialmetaphern der Verteilten Künstlichen Intelligenz beschäftigen sollte. Zeitschrift für Soziologie, 26 (1):3-22, 1997.
- [18] M. Matheas. An Oz-Centric Review of Interactive Drama and Believable Agents. Technical Report CMU-CS-97-156. School of Computer Science, Carnegie Mellon University, Pittsburgh, PA., 1997.
- [19] H. J. Müller, T. Malsch, I. Schulz-Schaeffer. Socionics: Introduction and Potential. *Journal of Artificial Societies and Social Simulation*, 1 (3) <http://www.soc.surrey.ac.uk/JASS/1/3/5.html>, 1998.
- [20] [A. Schefflyzyk]. Moscow-Tartu School. In T. A. Sebeok (ed.), *Encyclopedic Dictionary of Semiotics*, vol. 1, pp. 166-168. Mouton de Gruyter, Berlin, New York and Amsterdam, 1986.
- [21] P. Sengers. Practices for Machine Culture: A Case Study of Integrating Cultural Theory and Artificial Intelligence. *Surfaces*, 8, 1999.
- [22] W. I. Thomas. *The Unadjusted Girl*. Little, Brown and Co., Boston, 1923.
- [23] P. Torop. Cultural Semiotics and Culture. Sign Systems Studies, 27: 9-23, 1999.
- [24]U. Volli. Manuale di semiotica. Laterza, Roma-Bari, 2000.

C
POSTER
S
S
I
G
N
2
0
0
1

Interpoetry: a game of words, images and sounds as a poetic sign in digital media

Jorge Luiz Antonio (reviewer) Pontifical Catholic University of São Paulo, Brazil, PhD candidate Rua Goiânia, 7, Alto das Palmeiras,

CEP 13.301.341, Itu, São Paulo, Brasil 055 011 4023-2704 jlantonio@uol.com.br Philadelpho Menezes (1960-2000) (cd rom's co-author) Pontifical Catholic University of São Paulo, Brazil, professor

in memoriam

Wilton Azevedo (cd-rom's co-author) Presbiterian University Mackenzie, Brazil, professor Avenida Angélica, 546, apto. 43, Higienópolis, CEP 01228.000, São Paulo, Brazil 055 011 3662-3690

wilazevedo@sti.com.br

ABSTRACT

This paper treats about the hypertextuality and interactivity of Philadelpho Menezes and Wilton Azevedo's cd-rom *Interpoesia* (*Interpoetry*).

By describing the procedures of poetic construction, it is analyzed the poetic sign as a game of words, images and sounds in digital environment.

The prefix *inter* plays an important role in the cd-rom project not only as a new term for digital poetry, but also as a cultural and technological product that intends to replace the book.

Keywords

Poetry and Technology - Digital Poetry - Interpoetry

1. INTRODUCTION

Interpoesia: Poesia Hipermídia Interativa (Interpoetry: interactive hypermedia poetry) is a cd-rom created by Philadelpho Menezes (1960-2000) and Wilton Azevedo, in 1997/1998.

Interpoetry is the result of a project integrated with digital supports, and realized by Estúdio de Poesia Experimental (Experimental Poetry Studio) of the Post-Graduation Program in Communication and Semiotics in PUC - SP (Pontifical University of São Paulo) and by Post-Graduation Program in Education, Art and Culture History of the Mackenzie Presbyterian University (São Paulo).

First published at COSIGN-2001, 10-12 September 2001, CWI, The Netherlands.

2. INTERPOETRY: AN INNOVATIVE PROJECT

It is an innovative project made by university professors which participate of several electronic art exhibitions.

Menezes and Azevedo have created a new term and concept for a special type of digital poetry: *interpoetry*, or *interactive hypermedia poetry*: "poems in which sounds, images and words coalesce, in a complex intersemiotic process, in a technological environment which precisely facilitated the simultaneous presence of verbal, visual and acoustic signs: hypermedia programs" (Menezes in Menezes e Azevedo 1997/1998), that is, "an intersign exercise which makes clear the significance of the sign traffic or digital media, bringing about what could be called a new era of reading (Azevedo in Menezes e Azevedo 1997/1998).

The poetic synthesis proposed by the authors intends to bring together the verbal, sound and visual poetry in a context in which the interactivity overreaches the concept of intertextuality, once the dialogue with other works of art and authors realized in the electronic and digital environment, totally suitable to it, neither transposed, nor adapted.

It is a work of digital interpoetry that makes a dialogue with other types of text, for the "fusion of genre is, furthermore, natural to interpoetry: visual peotry, sound poetry, theoretical text, encyclopedic information, fiction, lies, games, all are possible paths within the interpoem, including the possibility of entering into commerce (or dialogue) with non-technological media" (Menezes in Azevedo e Menezes 1997/1998).

3. CONCLUSION

The use of several resources from many vanguard movements in 20th Century, besides the dialogue between various arts, leads us to observe, in this cd-rom, a new poetic language, that results of the fusion of verbal, sound and visual poetries in the electronic-digital environment, what makes the work a cd-rom of *digital interpoetry*.

Another aspect of the cd-rom to be pointed out is: as it is not possible to take a copy in paper, the reader-operator needs to access the cd-rom several times and, by this way, she/he makes a different reading based in visualizing and clicking mouse to get different meanings each time, because her/his memory makes different associations. The bricolage effect of the each interpoem leads the readeroperator to do the same in her/his mind.

The expression *interactive hypermedia poetry* corresponds to a group of concepts with which the digital interpoetry is composed: it is *poetry*, art of words and sounds, in the oral and written context that leads to visual; it is also *visual poetry*, that is, spatial forms of the word, the geometric form given to it, or the ideogramatic, pictorial and visual signs; it is *hypermedia*, with the meaning this word has for technological poetics, as multimedia, that is, the immense variety of ways of communication in interconnection; and *interactive* as a communication system between computer and human, in their various interfaces.

Taking Pierre Lévy's statement, we can affirm *Interpoesia: Poesia Hipermídia Interativa* contains the cyberculture art features: the participation of whom experience, interpret, explore or read it, and bears the typical organization of a collective creation, the continuos creation (Lévy 1997: 94-95), the interactivity between reader-operator and poetoperator-technician, the interface between man-readeroperator, between the microcomputer and the mouse, an inter-hyper-textuality that takes us to another texts and arts, in an complex intersign tissue, which characterizes the postmodern culture.

Along with other authors who dedicate to digital poetry, this cd-rom assures its presence as an unique and special example of the Brazilian digital poetry, justifying also the title of *interpoetry*, for it indicates the new ways of poetic experimentation and research in the new technological media.

4. EXHIBITIONS

The cd-rom was exhibited in the Cultural Space João Calvino, at the Mackenzie Presbyterian University, in São Paulo, from May 29th to June 6th, 2000, and it is also part of especial exhibition *Art and Technology* (section of Cyberart: zones of interactions, Hypermedia: CD-ROMs) in II Biennial of Visual Arts of Mercosur in 1999/2000, in Porto Alegre, Rio Grande do Sul, Brazil, and in *13th Brazilian Symposium on Computer Graphics and Image Processing*, in October 2000, in Gramado, Rio Grande do Sul, Brazil.

5. ACKNOWLEDGMENTS

My thanks to ACM SIGCHI for allowing us to modify templates they had developed.

A special thanks for the Professor Wilton Azevedo's help during the preparation of this paper.

6. REFERENCES

[1] Antonio, J. L.

2001 - Interpoesia: poesia hipermídia interativa. In:

http://www.otal.um.edu/~rcss/books/menezes.html

- [2] Bairon, S. (1995). Multimídia. SP, Global. (Contato Imediato).
- [3] Lévy, P. (1997). Quatro obras típicas da cibercultura: Shaw, Fujihata, Davies. In: Domingues, d. (org.) A arte do século XXI: a humanização das tecnologias. Trad. div. SP, Ed. Unesp, (Primas), p.95-102.
- [4] Menezes, P.

1988 - Demolições (ou poemas aritméticos): 1983-1986. SP, Arte Pau-Brasil.

1991 - Poética e visualidade: uma trajetória da poesia brasileira. Campinas, SP, Ed. Unicamp. (Viagens da Voz).

(org.) 1992 - Poesia sonora: poéticas experimentais da voz no século XX. SP, Educ. (Caleidoscópio).

1994 - A crise do passado: modernidade, vanguarda, metamodernidade. SP, Experimento.

1996 - Poesia visual: reciclagem e inovação. Em *Imagens*, revista, Unicamp, janeiro/abril 1996, p. 39-48.

1996a - Poesia sonora: do fonetismo às poéticas contemporâneas da voz. SP, cd rom, Laboratório de Linguagens Sonoras do Programa de Pós-Graduação em Comunicação e Semiótica, PUC SP / Microservice / FAPESP.

1997 - A oralidade no experimentalismo poético brasileiro. In: Domingues, Diana (org.). A arte no século XXI. SP, Ed. Unesp, (Primas), p.2722-281.

1998 - From Visual to Sound Poetry: the technologizing of the word, in *Face*, revista on-line, SP, SP, PUC-SP/COS, 1° semestre 1998, in <u>http://www.pucsp.br/~cos-</u>puc/face/s1_1998/poesia2.htm

[5] SCOTT, D.

1988 - Pictorialist poetics: poetry and the visual arts in nineteenth-century France. Cambridge / New York / New Rochelle / Melbourne / Sidney, Cambridge University Press. (Cambridge Studies in French).

1997 - CD traz experimentos com poesia sonora. In *O Estado de S. Paulo*, jornal, 15.03.1997, Caderno Cultura, p. D9.

[6] ACM SIG PROCEEDINGS template. http://www.acm.org/sigs/pubs/proceed/template.html.

7. APPENDIX

Technical Specifications

Minimal configuration: Pentium 166, 16MB

Instructions to play the Cd-Rom

To execute the Cd-Rom, please open the directory "Interpoesia" and access the icon "Interpoesia.exe"

Some instructions should be presented in order to appreciate the cd-rom better. It intends to be a guide for English speaking people. Here they are:

- first of all: with the key "esc" you can stop playing the cd-rom; then, you need to start again

- getting the cover, on the left, you click with the mouse the letter "T" (read and white), for translating into English

- as soon as you press "T" (translate), put the mouse indicator on the names of *philadelpho menezes* or *wilton azevedo*, in order to read authors's manifestos on interpoetry

- the interpoems are presented in two columns, one named "i" and other named "p", as follows:

 "i": O lance secreto (The Secret Move) Reviver (To revive) O inimigo (The enemy) Máquina (Machine)
"p": Missa (Mass) Lábios (Lips) Somatória (Adding) O Tigre (The tiger) Vírus (Virus)

Credits

Poems by Philadelpho Menezes and Wilton Azevedo

Poems by Philadelpho Menezes by means of Ana Aly's final art

Art Direction by Wilton Azevedo

Sound edition by Alessandra Vilela and Sérgio Bairon

Translation into English by David Scott

Programming / animation by Alessandra Vilela

Compact disc 065.462 manufatured by Pólo Industrial de Manaus by Sonopress -Rimo da Amazônia Ind. e Com. Fonográfica Ltda, in Manaus, Amazonas, Brazil.

Narratives Argument for Interaction Models: Or how our unstable double-agent found its way into the textual machine

Jaap Drooglever Haarlemmerstraat 243 2312 DR Leiden Netherlands JAAPDR@DDS.NL

ABSTRACT

In this paper a semiotic reading of Andruid Kerne's CollageMachine will result in the conflation of the software agent with the narrative agent of post-structural semiotics. This reader orientated approach can lead to the exploration of new and exiting territory

Keywords

Software agents, narrative agents, CollageMachine, reader oriented agents.

1. POST-STRUCTURAL INTRODUCTION

'He did it without a sound,' said one. 'He cut their throats before they even knew it.' That was not strictly true, since the wife and children had seen me on the stairs, but, still, it is the thought that counts. 'E must be invisible,' a woman was whispering to her neighbour. 'Nobody saw 'im come or go.' I wanted to thank her for her flattering report but, of course, I was compelled to be invisible among them again. 'Tell me,' I asked an odd-looking fellow with a red scarf knotted around his head, 'was there much blood?' [1]

More secret than this *agent*, it cannot be. Spying over the ones gossiping about the murder he himself committed. He actually participates in the creation of stories about his own cruelties. Committed in this situation, committing in another he is the perfect *double-agent*, and we will need a real Holmes to solve this crime. In the following paper will be tried to elaborate the different meanings of *agent* in software- and narrative theory.

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

Not to solve anything, that I will leave to the good people. In due conformity with the rules of science, will tried to make the problem slightly more complex. This will be done by giving a go at defining a type of *agent* combining the best of both, resulting in the definition of an *agent* for the application of an interactive narrative, making use of the possibilities of the new and so-called unstable media. This new type of *agent* can be best described as an unstable *double-agent* of interactive narrative.

In doing this, there will be made use of the semiotic theory of narrative as is developed by Mieke Bal and others. In the broad field of semiotic and narrative theory, the choice is for her theories since they are very precise and they allow for an approach to the technical laws of narrative in a way that very effectively combines the power of textual structures and the role of the reader in the process of making meaning out of texts. In the view of Bal, the word 'text' is not only confined to language as such but it can be applied to other cultural products like images, paintings and architecture as well. It is the reader who interprets these phenomena, making use of the text and her or his own judgements about it. Therefore every interpreting act is an interactive process. This reader orientated approach, which does not lose itself in absolute relativism of interpretation, offers a good framework for this paper and this conference as well.

One general movement of post-structural semiotics is the movement away from classification and representation towards a productive form of theory: not so much the epistemological status of the theory is what counts, but more the things you can do with it: how much knowledge can be gained, what constructions can be made or found. The quality of the theory is measured by its productivity. Meaning in this view is a product of the a process between reader and text, this means the end of a theory orientated towards representation and the beginning of a theory where difference is the key notion. As such there will also be made us of the theories of Charles Sander Peirce and Gilles Deleuze. In my view this theoretical basis of Bal, Peirce and Deleuze offers the greatest possibilities for a semiotic theory of interaction. And such a model is needed, for although DVD, Dolby Surround and such show a definite richness of media, they also too often show an equally poor understanding of the process of the production of meaning, which is prevalent in the common bombardment of signs the reader of these media is exposed to. A semiotic fact is that subtle signs can produce more meaning than an overload of information. Such an overload kills the possibilities of interaction in an environment where much information is not fixed but real-time producible.

In the tradition of Bal the definition of a narrative is as follows:

"In a narrative an agent is narrating, by means of which activity this agent is giving his or her vision of an ordered sequence of actions leading to events." [2]

This definition accounts for traditional narratives. These are fixed texts. As is mentioned above, every interpreting act is an interactive process. But still there are differences between a text which makes use of the possibilities that the new media offer and a traditional fixed text. What is added in an interactive narrative is an unstableness that makes it possible for the reader not only to have a role in the interpreting process, but also in the organisation of the text itself. In an interactive narrative, the reader can have a role in defining both the context and the route taken through the text. In terms of Peirce, the reader gains an extended role on the level of thirdness. This is the level where object shifts into meaning. For example a red traffic light: the level of secondness would involve 'object = stopping' while the level of thirdness would be the actual activity the reader undertakes as the result of percepting the red light sign. And this could just as well be: speeding up a little to make it in time to the other side of the street. This active result is the most important movement of the level of thirdness. The extension of the role of the reader on the level of thirdness is a decision not only involving the meaning of the sign the reader has read but also a decision about which signs will be read in the future. The unstableness of the text deepens the role of the reader. Whereby a reader of a static text has a role at the level of secondness choosing iconic and indexical relations and at the level of thirdness choosing iconic, indexical and symbolic relations, the reader of an unstable text also has an involvement at the level of firstness in the selection process of representamens.

2. THIS DOUBLE-AGENT

To show how theoretical and practical consequences conflate, underneath an example will be worked out. The application discussed here is called *CollageMachine* and is made by Andruid Kerne, who is visiting professor in computer science/multimedia at Tufts university and received support for this project by the NYU Media Research Lab. In an article in *Leonardo*, *CollageMachine* is described as follows:

> "CollageMachine builds interactive collages from the Web. First you choose a direction. Then *CollageMachine* will take you surfing out across the Internet as far as it can reach. It builds a collage from the most interesting media it can find for you. [...] *CollageMachine* deconstructs Web sites and re-presents them in collage form. The program crawls the Web, downloading sites. It breaks each page down into media elements – images and texts. Over time, these elements stream into a collage. Point, click, drag and drop to rearrange the media. How you organize the

elements shows *CollageMachine* what you're interested in. You can teach it to bring media of interest to you. On the basis of interactions, the evolving model informs ongoing choices of selection and placement." [3]

Another term used in this article is "An Interactive Agent of Web Recombination" besides the fact that the first description is a lot more smooth, this last one contains some interesting triggers for narrative theory. The definition of narrative is strongly interwoven with the notion of *agent*. Without this *agent* there is no story. It is defined as "a textual subject that produces the narrative". [4] Another important feature is the 'actions leading to events'. Both features point at a definite subjectivity that is built in to the narrative mode. The source of this subjectivity is the *agent* and it becomes manifest in the way this *agent* orders actions into events. This subjectivity is one of the most productive factors of the narrative mode, it can give shape to suspense, desire, curiosity and what not. This subjectivity is the factor that makes of breaks the story. The subjective telling of the tale is what makes it interesting for the reader to read and the subjective ordering is the added value of plain enumerating opposed to the narrative mode. Namely a sense of subjective direction, from actions to events. In this subjectivity lie the greatest possibilities but also the greatest threats to interaction. At the one hand, if a new media project wants to employ the narrative mode, it should not be afraid of the subjective agent, for it is constitutive to the narrative mode. On the other hand: a subjectivity located only on the side of the text, kills any possibility for interaction, leaving no room for interactive new media applications whatsoever. Semiotic theory made clear that this form of a subjectivity residing only at the side of the text, is fictional but the definitions of narrative, and the agent mentioned above are tied to a static text, not accounting for the differences between static and interactive texts. In this paradox the CollageMachine gives a brilliant solution.

In software the agent is, of course, a much debated operator. Using it as a keyword in Google alone results in nearly one and a half million of hits. This enormous amount of debate accounts for a set of definitions that is nearly as diverse. Examples of definitions are. "Agents are active and ever-present software components that perceive, appear to reason, act and communicate". [5] Is one, another is: "An intelligent agent is a program that deploys own initiatives.". [6] Both definitions overlap each other and could well be about the same thing, as is the case. Yet, that is not necessary. With the software agent we are dealing with an abstract phenomenon that is as vague as abstract thought can be. This abstract element 'agent' only acquires a specific meaning in an application of it. And I am happy to be able to announce, that the definition of the software agent in the application CollageMachine, given by Andruid Kerne fits quite well the *agent* in semiotic narrative theory. Andruid Kerne:

> "By an agent, I mean a program that acts on behalf of the user. The agent makes decisions on its own volition. It learns about the user's interests. It can run autonomously without direct

input from the user. It adapts in response to the user's ongoing expression. One type of agent is called a *recommender system*. These offer suggestions to the user about interesting content. Another type creates graphic representations based on a set of constraints. *CollageMachine* is both of these." [7]

The agent of the CollageMachine makes choices on the basis of a set of restraints, at the same time it recommends. How can that be done? Seen from the point of view of a static text, recommendation should take the form of the question, accept or reject this part of information into the text. This would mean a standard hypertext, where the reader can choose which link to click or a system where the reader gets the one question accept/reject after the other. In such a system, the subjectivity of the agent is lost and transferred to the reader, who is the one choosing actions. In such a model we are not dealing with a narrative: for hypertext gives choice but not direction, there is no ordering of actions leading to events, and there is a danger that the text stays only on the level of sole actions in their own context, producing no surplus of meaning. In the CollageMachine recommendation takes another shape, based on an interactive conception of a text. The choices made by the reader are taken in the process of a text continuingly changing. In an unstable environment presentation can also be interpreted as an offer to choose this or that direction for future selections in the process. Seen from the unstable point of view recommendation can take the shape of presenting for a reaction to the reader, in stead of a question accept/reject admission into the presentation. In this model the reader does not choose the actual presented content but has a role in the direction the agent is heading.

This model of the *agent* can be best described here with the term *double-agent*. This doubleness does not only account for the fact that this software *agent* can also be seen as an *agent* as is described in narrative theory but its doubleness is most inventive and powerful because the model of the *agent* in the *CollageMachine* implies a shared subjectivity. On the one hand, the *agent* keeps its crucial role in ordering actions into events, on the other hand the reader has a defined role in the process of choosing which actions will lead to what event. The *agent* is the leading subjectivity in the application, as is required for an authentic narrative but at the same time the subjectivity of the *agent* is influenced up to a large part, by the reader.

3. MOVING ON SUBTLY...

In *CollageMachine* this *double-agent* has taken one form but so many others can also take shape, maybe making a better use of the examples that the already existent tradition of thousands of years of narrative brings us. Above is mentioned how many of the existent new media products like DVD and Dolby Surround combine a richness of media with a poor understanding of the process of the production of meaning. This could be labelled as the *emotion-winpoints effect*. Such a misconception is prevalent in most digital games. In order to win you have to score points, which are visualised very obtrusive on the screen so the player is obliged to identify with them. The player of such a game does not

identify her or himself with a character, but with the meters and the emotion-winpoints semi-representing it.

Two misconceptions lie at the basis of this emotion-winpoints effect. The first it shares with the CollageMachine namely mistaking the *agent* for a mastermind. This can also be found in old-fashioned 19th century novels: one auctorial *agent* having a hand in everything that happens. Another example is the fight and win plot. But what modern masterpiece still works with such base conceptions? In the narrative theory of Mieke Bal, different types of agents are analysed. Two larger categories are the focaliser and the actor. They all play with the implications of subjectivity. Subjective vision, underdog positions, local differences or different relations an agent has with other elements in the narrative. One narrative, in fact, can have many different agents, all having their own subjectivity making up the whole of the narrative together. The aim is not to mastermind the total but to play a constitutive subjective role in the interactions between al the different *agents* that make up the subtlety of the text.

The second misconception involved in the emotion-winpoints effect has its basis in the attempt to mirror the reader and her or his intentions in the characters of the game. In a semiotic conception this is the wrong thing to do. It is what above is aimed at by stating that we have to move away from representation since the sharing of subjectivity in the doubleagent must not result in an attempt to replicate the reader and her or his intentions but in a lively interaction between different subjective responses. When a reader indicates that she or he likes a certain situation, must the agent then respond accordingly? It could well be better, to let the *agent* make use of this piece of information in the further proceedings of the story. This can be done in many ways, and certainly not only in the form of an exact copy of the readers response in the agents subjectivity. Theoretically this is not even possible. What is possible, and what is much more interesting as well is an ongoing semiosis between the subjectivities residing at the digital end and the subjectivities residing at the living end of our interactive meaning making process.

In order to overcome the deficiencies and to make use of the chances a semiotic approach to new media, we'll have to let our imagination explode into all the 32 million colours that lie at hand. We don't have to search in the rich narrative tradition for long, to come to the conclusion that a good narrative actually has the power to generate strong emotions in the reader. Whether it be a good children's book, a piece of literature or a good film. For isn't it so that only the subtlety of Hannibal Lecter can conjure up the physically felt abjection that so many readers of Silence of the Lambs or Hannibal felt while sitting appalled in their movie theatre chairs? And now we're talking about cutting up bodies again, our *double-agent* offers opportunities not only to text or images but has an inspiration in the whole body and its faculties. What about tactile media, heartbeat response, response to music, sudden movements and what not? Whether a lie detector is really 100% safe or not, deploying it for an interactive narrative, not caring one damn about truth, could deliver a beautifully intense experience.

4. REFERENCES

[1] Peter Ackroyd: "Dan Lemo & The Limehouse Golem". Vintage: 1998 (1994) pp. 192

[2] Frans-Willem Korsten: "The Wisdom Brokers, Narrative's Interaction with Arguments in Cultural Critical Texts." Amsterdam: ASCA, 1998. pp. 13

[3] Andruid Kerne: "<u>CollageMachine, An Interactive Agent of</u> <u>Web Recombination.</u>" In: *Leonardo* Vol. 33 No. 5 pp. 347-350 (2000)

[4] Frans-Willem Korsten: "The Wisdom Brokers, Narrative's Interaction with Arguments in Cultural Critical Texts." Amsterdam: ASCA, 1998. pp. 13

[5]Gary A. Berg: "Human Computer Interaction (HCI) in Educational Environments: Implications of Understanding <u>Computers as Media</u>". In: *Journal of Educational Multimedia and Hypermedia* (2000) 9(4). pp 349-370

[6] http://www.cs.vu.nl/vakgroepen/cs/term_projects/agents.html -own translation

[7]Andruid Kerne: "<u>CollageMachine, An Interactive Agent of</u> Web Recombination." In: *Leonardo* Vol. 33 No. 5 pp. 347-350 (2000)

* Mieke Bal: "Narratology, Introduction to the Theory of Narrative." 2nd edition. Toronto, Buffalo, London: University of Toronto Press, 1997

* Andruid Kerne: *CollageMachine* (artwork): http://mrl.nyu.edu/ecology/collageMachine.

SENTIENT VR ENVIRONMENT DESIGN: THE MEMESIS PROJECT

Dr. Alison McMahan, M.A. Program Director, Film and Television Studies

The University of Amsterdam, Nieuwe Doelenstraat 16, 1012 CP Amsterdam The Netherlands

drajm@compuserve.com

ABSTRACT

The purpose of the *Memesis* project is to design an interface and navigation system that "reads" input off the player/user (using technologies derived from systems used for polygraph and blood pressure testing, among others) in order to provide the user with a tailor-made experience based on sub-conscious as well as conscious input. The program will be designed first for use in CAVEs (computer automated virtual environments), and then for online and console playing. For the project prototype I intend to develop a "horror/thriller/suspense" experience.

Keywords

Memesis project, interactive narrative, horror/thriller/suspense genre, interface design, cave.

1. INTRODUCTION

The purpose of the Memesis project is to design an interface and navigation system that "reads" input off the player/user (using technologies derived from systems used for polygraph and blood pressure testing, among others) in order to provide the user with a tailor-made experience based on sub-conscious as well as conscious input. The program will be designed first for use in CAVEs (computer automated virtual environments), and then for online and console playing. For the project prototype I intend to develop a "horror/thriller/suspense" experience. I will apply the knowledge gained from years of teaching screenwriting and cinema studies to dissect the horror genre as we know it from the cinema and translate it into a modularized version that players can enjoy in a 3D environment. Enjoyment of the environment will come from the direct interaction between user and environment (the user makes conscious choices) and indirect interaction.

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

The plan is to develop the project in cooperation with the CAVE at SARA (part of the University of Amsterdam in the Netherlands), the Alterface Company in Brussels, and Jeffrey Shaw at the ZKM in Kalsruhe. The project has just been awarded the support of a Carnegie-Mellon Fellowship.

The first phase of the *Memesis Project* is to create a horror-genre experience in a fully immersive interactive environment. There are three elements to this task: a new approach to interface design; a new way of thinking about and designing the relationship between the user and the interactive fiction; a modularization approach to the horror/thriller genre as we have come to know it from cinema, television and theatre.

2. BASIC ASSUMPTIONS

The *Memesis Project* starts off with the basic assumption that users seek out immersive digital experiences because they are looking for experiences that go beyond spectacle. *Memesis* will apply Daniel Dayan's and Elihu Katz's distinction of 'ceremony', 'festival' from 'spectacle'. In contrast to spectacle, ceremony and festival involve interaction, although to different degrees: festival involves a strong interaction between audience and the event they experience (as at carnivals), whereas ceremonies involve a measured interaction, (such as shouting and applause). Film is firmly rooted within the society of the spectacle. Interactive digital media, by contrast, need to be conceived along the lines of festivals and ceremonies. We expect that the endusers will be attracted to the *Memesis Project* because it will allow them to explore their inner fears and anxieties and help them to learn more about themselves.

This requires a new approach to interface design. The *Memesis Project* aims for a greater degree of transparency and immediacy, goals of all new media technologies. However, we plan to achieve these goals in a different way. The relationship between input (how the user communicates with the computers system) and output (how the computer system responds to the user) is the first area that will change.

Part of this approach requires changing the relationship between diegetic (elements of the fiction in question that are perceptibly part of the story world) and non-diegetic content (elements of the fiction that are clearly not a part of the story world). The aim here is to make the *Memesis* experience as transparent as

possible, but also to shift away from the computer-game approach which takes the important presence of non-diegetic interaction for granted.

3. HUMAN COMPUTER INTERFACE DESIGN

The nature of the *Memesis* experience necessarily entails a way for the game software to get readings off the player without the player having to constantly make conscious choices within the game. This can be achieved in a variety of ways. First-phase research will focus on adapting polygraph device readers to the *Memesis* interface. Second-phase design will include galvanic skin readers and eye scanners as well pressure-sensitive floor mats. The key is to take these readings with the least amount of discomfort and annoyance to the user. In its initial phase the project is designed for use in CAVEs, but we hope to eventually develop PC and/or console versions. Ergonomic research into these devices will be done with this eventual goal in mind.

4. USER-PROGRAMME RELATIONSHIP

Barbara Becker has noted that people are driven to create new identities for themselves on-line, in MUDs (multi-user domains such as text-based adventures like *Genesis* or 2D graphic MUDs like *Ultima Online*) because they want to "explore new dimensions of identity by its virtual (re-) construction" and in order to become "the product of ones own tinkering."

In MUDs and MOOs users work towards such goals by creating avatars (textual or graphic representations of themselves that include a character designed to fit into the fictional environment in question, complete with a set of personality traits, skills, and health status). Memesis enables users to interact with the environment "as themselves", rather than through the use of an avatar. What this means in practice is that user is not assumed to be a "co-author", as most theorists of new media, including Barbara Becker and George Landow, believe. Even in MUDs and MOOs where a maximum degree of control has been provided to expert users who achieve the status of "wizards," these controls are non-diegetically based. In other words, such expert users then abandon their status as players almost completely and limit their participation to regulating the MUD or MOO environment "from above", i.e., from a non-player position. In fact, one of the requirements of wizard status in a MUD is that the player, in addition to having a relatively high degree of programming skills, must have completed all of the quests and adventures available in the MUD.

The design for *Memesis* starts with a different assumption: rather than assume that the user is a co-author, we believe that part of the appeal of interactive fictions is that they enable a different level of narratee engagement. (Here I am following Edward Branigan, and his use of these terms in *Narrative Comprehension and Film*). Historically, certain classes of readers have chafed under the limitations of narratee construction, such as the fans of long-term popular shows like *Star Trek* who engage in story-tree making. Interactive fiction enables users to tailor their position as narratee to suit their own interests, though this "tailoring" is still limited by the parameters of the program design. *Memesis* is designed to capitalize on this heretofore unrecognized aspect of interactive fiction. This aspect of the design recognizes virtual reality environments as the ultimate development of televisuality. Of interest here is Elihu and Dayan's semiotic theory of ceremony, festival and carnival. Elihu and Dayan applied these concepts to televisual events such as Lady Diana's wedding, but these concepts can also be applied the altered relationship between narrator and narratee positions brought about by the modularization of the narrative, and it is one avenue we intend to explore in *Memesis*.

5. THE MODULAR HORROR EXPERIENCE

Memesis, in its first incarnation, will consist of a modularization approach to the horror/thriller genre as we have come to know it from cinema, television and theatre. The horror genre can be subdivided into various types, such as "shock horror", "body horror", "psychological horror" and "moral or psychic horror." Shock horror depends on certain effects that are based on our fear of being attacked. Most shock horror works by activating viewer's phobic responses. VR environments are often used by psychologists to help phobic patients (patients can learn how to cope with vertigo or agoraphobia in a VR environment, for example). Memesis builds on such experiments to incorporate shock horror into the fictional experience. Body horror works by activating the revulsion/repulsion response in viewers. Certain VR techniques have been developed for use in surgery, etc, and Memesis plans to take advantage of the special qualities of VR vision to elicit similar responses in users of the Memesis fictions. Psychological horror works by giving viewers a particular insight into a horrible situation or a horrible mind, such as the mind of a psychopath, and even allow us to identify with him. Memesis will work by enabling the user to explore the darker corners of their own mind. This is the most challenging design element for Memesis, as this type of horror depends on a high degree of control from the narrator position.

However, for precisely the same reason moral or emotional horror can be more effective in the VR environment, because the altered status of the narratee implicates the user morally and emotionally into the situation. The hypothesis we wish to test in *Memesis* is that by enabling the user to provide the computer with input at a subconscious level, the resulting interactive experience will be profoundly more effective at the moral and psychic level. This is why, of all the storytelling genres that would lend themselves to this experiment, the horror genre was chosen: because moral and psychic effect is precisely what the horror genre achieves best. Future versions of *Memesis* are planned for other genres, such as fairy tales and romantic comedy.

The first step towards creating a modular horror experience is to analyze the horror genre itself, as it has manifested so far in film, television and prose fiction. Once an initial survey of the literature is complete, typical horror experiences can be precisely defined and broken down into units or modules of experience that *Memesis* can recombine according to player input.

6. THE A.I. OF MEMESIS

This brings us to the role of artificial intelligence (AI) in *Memesis*. The AI in *Memesis* will manifest in two ways: in the

intelligence of the environment and in the intelligence of the bots (characters generated by the program).

As currently envisioned, *Memesis* will consist of a series of individual experiences, each about 15 minutes long (the average user cannot spend more time in a 3D environment). The first part of a typical 15 minute experience will consist of an entry experience designed to gather readings from the user while the user is entertained with instances of shock horror. Then there will be the introduction of a bot which will continue to gather readings from the user. This gathering will work primarily on the indirect level even though user and bot will also have a direct engagement.

The bots in Memesis will not be designed as synthespians, or synthetic actors, such as the bot ELIZA designed by Joseph Weizenbaum. The Memeis approach is closer to that of Stephen Grand, the creator of the computer game Creatures. Grand created Norns, the creatures that inhabit Albia, the virtual world of the Creatures series, first released in 1996. Players rear their norn and help it through various encounters. Grand programmed his norns with over 300 genes and simple drives to satisfy, such as the drives to eat, breed, and avoid pain. Most importantly, norns can learn by doing. Productive skills are reinforced, unused ones fade away. This kind of design characterizes most virtual pets, but Grand also designed a "biochemistry" for his norns that behave much as hormones and neurotansmitters do in the human body. Players can influence the norns' learning pattern by doling out rewards and punishments, and when norns breed, the "DNA' patterns are mixed, often with unexpected results.

The bots in *Memesis* will also learn from experience, by incorporating information gathered directly and indirectly from a series of players. This information-gathering mechanism will be based on psychological testing instruments that alert clinicians to phobias and deep seated fears. The *Memesis* bots will not simulate human adults. They will not, for example, in the first phase at least, be capable of synthetic speech. Instead they will follow the limited rules of their environment. Their ability to provide an impression of intelligence will be based on their ability to read the player and the way the current readings are added to and blended with experiences with past players. For example, interactions that feature a fear of insects in today's player will build on interactions with previous players that also had an insect phobia.

The main event in each visit to *Memesis* will be a singular interactive horror experience designed to evoke one of the emotions or moral reactions traditionally evoked by the horror genre, such as fear of abandonment, fear of betrayal, fear of loss, fear of humiliation, and concomittantly, the fear of being discovered at wrongdoing such as abandonment, betrayal, etc. For example, almost every adult is familiar with the fear that an infant put into their care comes to harm. The first *Memesis* experience to be built will be designed with this fear in mind.

After the main event there will be a brief "exit" experience designed to enable the user to adjust to the end of the visit to *Memesis*.

7. SUMMARY

The *Memesis* Project is an art project, still in the design phase, which will consist of an interactive horror experience with a unique human computer user interface, an interface that will read direct and indirect input from the user. These responses will then be used to provide the user with an interactive horror experience designed around the same limited number of phobias and deep-seated fears that are characteristic of the horror genre in other mediums, such as film. The overall goal is to entertain as well as to provide a space for self-reflection. What we learn from this experience can then be applied to developing other interactive VR experiences (using other genres), and as the basis for a new type of interactive fiction form.

For CV see

WWW.HUM.UVA.NL/~FTV/FACULTY/ALISON

8. RELEVANT LITERATURE

- Arseth, Espen, *Cybertexts: perspectives on Ergodic Literature*, Baltimore and London: The Johns Hopkins University Press, 1997.
- Bal, Mieke, *Narratology*, Second Edition, Toronto, Buffalo, London: University of Toronto Press, 1997.
- Branigan, Edward, *Narrative Comprehension and Film*, London and New York: Routledge, 1992.
- Chatman, Seymour, *Story and Discourse: Narrative Structure in Fiction and Film*, Ithaca and London: Cornell University Press, 1978.

Coming to Terms: The Rhetoric of Narrative in Fiction and Film, Ithaca and London, Cornell University Press, 1990.

Clover, Carol, Men, Women and Chainsaws: Gender in the Modern Horror Film, Princeton, New Jersey: Princeton University Press, 1992

- Damer, Bruce, AVATARS!, Berkeley, CA: Peachpit Press, 1998
- Davidson, Clive, "Agents from Albia" New Scientist, 9 May 1998 No. 2133, pp. 38-44.
- Dayan, Daniel, and Elihu Katz, "Electronic Ceremonies: Television Performs a Royal Wedding", in *On Signs*, edited by Marshall Blonsky, Baltimore, MD: The Johns Hopkins University Press, 1985, pp. 16-32.
- Dennet, Daniel C., *Brainchildren: Essays on Designing Minds*, London: Penguin Books, 1998.
- Duguet, Anne-Marie, and Heinrich Klotz and Peter Weibel, Jeffrey Shaw: A User's Manual, Cantz: Edition ZKM, 1997
- Garrand, Timothy, "Scripting narrative for interactive multimedia" Journal of Film and Video, Vo. 49, No. 1-2, Spring-Summer 1997, p.66-79.
- Haraway, Donna, Simians, Cyborgs, and Women: The Reinvention of Nature, New York: Routledge, 1991.
- Haugeland, John, Artificial Intelligence: The Very Idea, Cambridge, Mass: MIT Press, 1989 (1985).

- Hedges, Inez, Breaking the Frame: Film Language and the Experience of Limits, Boomington and Indianapolis: Indiana University Press, 1991.
- Heudin, Jean-Claude, ed., Virtual Worlds: Synthetic Universes, Digital Life, and Complexity, New England Complex Systems Institute Series on Complexity, Reading, Mass: Perseus Books, 1999.
- Landow, George P., Hypertext 2.0: The Convergence of Contemporary Critical Theory and Technology, Baltimore and London: The Johns Hopkins University Press, 1997 (1992).

ed., *Hypertext Theory*, Baltimore and London: The Johns Hopkins University Press, 1997.

Laurel, Brenda, ed., The Art of Human-Computer Interface Design, Reading, Mass.: Addison-Wesley Publishing Co., 1990

Computers as Theatre, Reading, Mass. Etc.: Addison-Wesley Publishing Co., 1993.

- Leonard, Andrew, *BOTS: The Origin of a New Species*, New York: Penguin Books, 1997
- Levy, Steven, Artificial Life: The Quest for a New Creation, London: Penguin, 1992.
- McMahan, Alison, "Spectator, Avatar, Golem, Bot: Interface and Subject Position in Interactive Fiction" paper given at the SCS 2000 in Chicago, USA.

"Social Stage, Simulacra, CivilzationTM" paper on citysimulation games given at SCREEN studies conference, Glasgow, July 2000.

"Verbal-Visual-Virtual: A MUDdy History" *Gramma: Journal of Theory and Criticism*, Aristotle University of Thessaloniki, Volume 7, 1999, pp. 73-90.

"The Effect of Multiform Narrative on Subjectivity" *Screen*, 40:2 Summer 1999 pp. 146-157.

"Interactive Narrative and the Interstitial Interface" *Interactive Frictions Conference*, The University of Southern California, June 4-6, 1999.

"Welkom in de speelautomatenhal. Het compuerspel als symptoom van het post-postmoderne tijdperk" (Welcome to the Arcade), pp. 45-57 and "Starring Lara Croft", pp. 213-222, in *Hollywood op straat: Film en televisie in de hedendaagse mediacultuur*, Thomas Elsaesser and Pepita Hesselberth, eds., Amsterdam: Vossiuspers, (AUP), 2000.

"De CAVE bij SARA (The CAVE at SARA)", Archis, 1:1999, p. 79-80.

"Ruimte als evenement: IMAX (Space as Event: IMAX)" Archis, 3:1999.

"Interactive ruimten:screenspace (Interactive Space:screenspace)" *Archis*, 4:1999, p. 78-80.

"Verhalende ruimte (Narrativized Space)", Archis, 6:1999, p. 78-80.

"Architectuur van de geest: Werk van Jeffrey Shaw" (Architecture of the Mind: The Work of Jeffrey Shaw") Feature Article, *Archis*, 6, 2000, p. 38-43.

"Interactive Lexicon" *Millimeter*, Teleproduction section July 1989 (analog interactive technology).

"E-Motional Control" *Millimeter*, Industry Report November 1989 (motion-control cinematography).

"BTV's Sharper Image" *Millimeter*, Teleproduction section September 1989 (high band corporate projection systems).

- Milano, Dominic, ed., Interactivity in Action: Case Studies of Multimedia Mastwerworks, Innovative Games & Other Successful Interactive Products, San Francisco: Miller Freeman Books, 1997.
- Morris, Richard, Artifical Worlds: Computers, Complexity, and the Riddle of Life, New York and London: Plenum Trade, 1999.
- Morse, Margaret, "Nature Morte: Landscape and Narrative in Virtual Environments," in Immersed in Technology: Art and Virtual Environments, edited by Dmary Anne Moser with Douglas MacLeod, Cambridge, Massachusetts: The MIT Press, 1996.
- Murray, Janet, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, New York: The Free Press, 1997.
- Propp, V. Morphology of the Folktale (Austin: University of Texas Press, 1968.
- Rabinow, Paul, ed., *The Foucault Reader*, New York, Pantheon Books, 1984.
- Rabinowitz, Lauren, "Introduction," pages 3-8, and
- Rugero Eugeni, "*Myst*: Multimedia Hypertexts and Film Semiotics," pages 9-26, both in *Iris* 25 (Spring 1998).
- Rimmon Kenan, Shlomith, Narrative Fiction: Contemporary Poetics, London and New York: Methuen, 1983
- Ryan, Marie-Laure, *Possible Worlds, Artificial Intelligence, and Narrative Theory*, Bloomington and Indianapolis: Indiana University Press, 1991.

, ed., Cyberspace Textuality: Computer Technology and

Literary Theory, Bloomington and Indianapolis: Indiana University Press, 1999.

"Interactive Drama: Narrativity in a Highly Interactive Environment", *Modern Fiction Studies*, Vol 43, no. 3, Fall 1997, p. 693-94.

- Samsel, Jon & Darryl Wimberly, *Writing for Interactive Media: The Complete Guide*, New York: Allworth Press, 1998.
- Sawyer, Brian & John Vourlis, "Screenwriting Structures for New Media", *Creative Screenwriting*, Summer 1995, pp. 95-103.
- Thalmann, Nadia Magnenat and Daniel Thalman, Artificial Life and Virtual Reality, Chicester, New York, etc.,: John Wiley & Sons, 1994.
- Turkle, Sherry, Life on the Screen: Identity in the Age of the Internet, London: Phoenix, 1997 (1996) (New York: Simon & Schuster, 1995).
- Tomkins, Silvan S. and Samuel Messick, eds., Computer Simulation of Personality: Frontier of Psychological Theory, New York and London: John Wiley and Sons, Inc., 1963.
- Darryl Wimberly & Jon Samsel, *Interactive Writer's Handbook*, Los Angeles: The Carronade Group, 1995.
- Varchol, Douglas J., *The Multimedia Scriptwriting Workshop*, San Francisco, etc: Sybex, 1996.

Metaphors and Multimodal Interaction

Dorothy Rachovides, Zoë Swiderski, Alan P. Parkes Computing Department, Lancaster University Lancaster. LA1 4YR. U.K. +44 1524 592326 {dot | zoe | app}@comp.lancs.ac.uk

ABSTRACT

Multimodality has the potential to facilitate richer interaction styles in both information retrieval and learning environments. However, its true potential will not be realised unless consideration is given to the application of *combined* modalities. This paper asserts that multimedia output from a system actually *requires* multimodality on the part of the user in order to ensure that the effectiveness of the communication or information is not lost. The notion of a "multi-modelling" approach to interaction along with the use of gestures and metaphors have been examined and two systems are described which attempt to implement these approaches.

Keywords

multimodality, interaction, multimedia, metaphors

1. INTRODUCTION

Information can be conveyed between people in a number of ways. People draw on a range of materials (e.g. pen and paper) and physical abilities (e.g. gesturing) in order to express themselves. While machines have become more prevalent as providers of information, the methods used to convey and receive information have undergone some radical changes. This evolution from human-human to machine-human information sharing was merely the beginning. It was not sufficient for machines to simply replace humans as providers of information; they had to provide larger quantities of well-presented information and more possibilities of interactivity.

The traditional keyboard/mouse interaction style has long been regarded as limiting in terms of expressiveness, efficiency and how naturally it can be used, which has led to an interest in the development of alternative methods of input. Similarly, the output produced by such systems has become more dynamic, and exploits enhanced graphical interfaces to provide an enriched

First published at COSIGN-2001, 10-12 September 2001, CWI,

The Netherlands.

audio-visual experience. Much development has occurred in terms of input and output facilities. Output, in particular, can consist of a rich structured and interlinked collection of multi-media objects. However, little thought has been devoted to the expressiveness of *combined* multiple input modalities. Users are unlikely to receive the information they require, or be able to refer to it appropriately when they receive it, unless the same consideration is given to the input as is increasingly being given to the output.

This paper discusses the relationship between multimodality input and multimedia output, with specific consideration of the implications of multimedia for the form, content and meaning of multimodality input. The potential for adopting a "multimodelling" approach to multimodality along with the use of metaphors is discussed. Finally, two systems which incorporate the use of metaphors are described.

2. THE IMPLICATIONS OF MULTIMEDIA OUTPUT FOR MULTIMODALITY INPUT

Over recent years, many systems have been developed for the dissemination of information. One example is a multimedia encyclopaedia. Another is the computer-based training system. In such circumstances, there is no longer a human information provider. The digital artifact assumes this role. This has consequences for the information receiver. Today, the information receiver, or user, is usually faced with multimedia information. This information may be delivered in a highly structured form, where the user is guided to the result of the interaction, so the interaction and control is probably limited. Alternatively, the information may be poorly structured, and therefore a high level of interaction and user control is subsequently required. Neither of these two extreme situations is optimal for the user. They may result in a confused user who feels overwhelmed or insufficiently informed, or is perhaps reduced to being little more than a "page turner" of an "electronic book". The output has a much higher level of expressiveness than that provided by the input.

The user has access to a wide range of modalities through which he or she typically interacts. Interaction should involve *bidirectional* communication. In human-computer interaction, while there is often great richness in the system's output, users' ability to make use of the modalities available to them is severely limited. Moreover, having received the multimedia output, the facilities available for the user to specify to the system what it is that he or she really requires are also typically very limited. Facilities such as pointing, clicking, and maybe simple speech and primitive gestures are provided for, but almost always each modality is considered in isolation from any other.

To date, most research in multimodal interaction has focused on using each modality separately, or in pairs, i.e. speech and gesture, gesture and gaze. Very little research has been carried out on the basis of a detailed analysis of human-human multimodal interaction. As a result, crucial contextual metadata are overlooked, such as the effect of facial expressions, gestures, or voice intonation on the meaning of an utterance.

However, the above discussion is not meant to imply that a multimodality interface will always be used in a multimodal way by its users. Oviatt [1] points out that human-human communication is typically a mixture of unimodal and multimodal interactions. In a well-designed system, individuals would be able to choose whether to interact multimodally or not. Often, this choice would be made on the basis of the activity being carried out, or the context in which that activity takes place. Though users favour the ability to interact multimodally, they do not always choose to do so, and they usually explore the use of each modality separately and then form their own pattern of interaction.

One problem in human-computer interaction is whether designers should tailor their systems to the user, or users should tailor their interaction patterns to the system. Considering multimodality interaction, the problem is the extent to which we can assume that multimodality will be exploited in a uniform way by different users. In Oviatt's study [1], users adopted either simultaneous or sequential integration patterns when combining speech and pen input. Each user's integration pattern was established early and remained consistent, but nevertheless, each user's pattern was unique. As an aside, it is probably the case that similar individual patterns of usage also apply to users' use of multimedia *output*.

A further important factor in multimodal systems is the extent to which the integration of modalities introduces *redundancy* in the content specified by different modalities. However, redundancy is often *complementarity*. The ability to convey the same information in several different modalities does not imply that a user will use all of these modalities to interact at any one time, but rather may choose which modality or combination of modalities is suitable at the given moment, in the particular context. Likewise, if the system produces output involving multiple media types, the user will often focus on a preferred media format, which may lead to the risk of missing important information. Thus, redundancy is often a useful property of multimedia output. The implications for multimodality user input has yet to be fully explored.

3. THE RELATIONSHIP BETWEEN MODALITIES

To us, the term "multimodal input" implies the existence of simultaneous or temporally co-ordinated expressions in a variety of modalities. The two most frequently combined modalities in human-human multimodal interaction, *speech* and *gesture*, are highly interdependent and synchronized during interaction. They are not always simultaneous, as gesture can often precede speech, or complement it by conveying information that is not explicitly uttered. Such cases typically involve a quick switch from speech to gesture and back to speech. This is accomplished so quickly and blended so naturally that it is perceived as simultaneous.

The view of linguists and some computer scientists that speech is a primary input mode has biased early multimodal systems towards speech input and "point-and-speak" systems. This has rendered speech to be the primary input mode in most multimodal systems in which it is included. Unfortunately, this has led to systems that consider other modalities that are employed as secondary, thus failing to recognize information that is not present in the speech. Speech is not the exclusive carrier of information. Even in a simple "point-and-speak" interface, it is possible to imagine a scenario in which both modalities in a particular activity are an indispensable component of the meaning of the "utterance". Consider telling a system to move a previously marked block of text to a new location:

> "move that" [spoken, accompanied by] pointing to block of text "to there" [spoken, accompanied by] pointing to target location

As this simple example demonstrates, when users interact multimodally they selectively eliminate linguistic complexities and replace them with an interaction pattern, which involves unimodal and multimodal aspects. However, what results is a complex "linguistic" structure in which meaning depends on the temporal and significant relationship between expressions in two modalities.

Different input modalities can be used to specify different content. The different modalities found in emerging technologies that recognise speech, handwriting, manual gesturing, head movement and gaze can significantly differ in the information they specify. They can also differ in their functionality during communication, the ways in which they are integrated with each other and their suitability for incorporation into different interface styles. In some cases, a given modality can be a simple analogue of another, in the sense that there is a direct translation between one and the other. However, in many cases, modalities vary in the degree to which they represent similar information, with some groups of modalities being more similar (speech and writing) than others (speech and facial expression).

4. TOWARDS "MULTI-MODELS" OF MULTIMODALITY INPUT

Multimodality has the potential to facilitate richer interaction styles in both information retrieval and learning environments. However, its true potential will not be realised unless consideration is given to the application of *combined* modalities, both simultaneously and over time. Progress has long been made in the structural and grammatical analysis of language, where the term is usually meant in the unimodal sense, as it applies to, say spoken English, of which the structural and semantic analysis is, of course a well-established field. However, mixed modality interaction, while drawing on the various languages of speech, gesture, etc., implies that account must be taken of the relationship between the simultaneously expressed statements from each of these languages. For example, the utterance "we'll get this paper finished by this evening", when accompanied by the quickly raised eyebrows of the speaker, might mean something quite different when accompanied by the speaker's reassuring smile. For multimodality interaction, then, the corresponding "grammar" would describe the structure of mixed modality "sentences", and the lexicon would map out the meaning of mixed modality "words". The meaning of an utterance would be inextricably linked with all of the multimodality components of the "utterance" and the relationship between them. In this respect,

what is required is a "*multi-model*" of multimodality communication. Such a model would enable us to specify and interpret mixed-modality inputs, and support an expressiveness and flexibility of input to match that increasingly found in forms of output.

At the time of writing, multimodality interaction in HCI is much less sophisticated than that offered by the combination of speech, gestures and other modalities found in everyday human-human interaction. However, even the standard typing, pointing, and clicking interface offers gestural possibilities (the selection of a portion of text with a mouse is essentially gestural, after all) that have hitherto been almost exclusively applied unimodally. Thus, the central argument of this paper applies to current, as well as future, systems.

Finally, we assert that multimedia output from a system actually requires multimodality on the part of the user. Communicating with a system about a diagram, for example, requires more than just speech, text and simple pointing. The effectiveness of a diagram may be lost if the participants in a discussion about that diagram must constantly translate their knowledge of the diagram into an alternative form to express it to the other participants. In other words, a final requirement of the "multi-model" of modality is that it considers the role played by the media that are referred to in by the input, since, for example, even the meaning of a simple gesture such as a wave of the hand will depend partly on properties of the *referent* of that gesture. The diversity of these properties of multimedia information will open up new expressive possibilities for multimodal communication in human-computer interaction. The "multi-model" of multimodality communication may provide a framework in which to address such issues.

5. METAPHORS

Metaphors are not a novel feature of HCI in themselves, the *desktop metaphor* being a prime example. However, the advent of multimedia and novel interaction techniques has perhaps overshadowed the effectiveness of metaphors with more focus directed at the core input and output techniques (e.g. speech and gesture input and multimedia output) rather than the underlying mechanisms, which will support them.

The advantages of metaphor usage in interfaces is not always apparent. One path of reasoning can be found in Umberto Eco's interpretation of the words of Aristotle "... the most ingenious and vigorous of Aristotle's conclusions, [is] that the metaphor is not only a means of delight but also, and above all, a tool of cognition." Eco also points out that Aristotle describes the creation of metaphors as " 'a sign of natural disposition of the mind' because knowing how to find good metaphors means perceiving or grasping the similarity of things between each other" ($to \ \delta\mu\omega\rho\omegav \ \theta\varepsilon\omega\rho\varepsiloniv$) (Poetics 1459 a6-8) [2]. Two projects are described in the following sections, which use metaphors extensively with a view to exploiting such similarities (both between media objects themselves as well as between real world actions and digital environments).

6. MULTIMODAL STORY CREATION -THE STORY CONDUCTOR

The story conductor under development by Dorothy Rachovides uses metaphors in two ways: (i) a series of visual metaphors are used to represent media types and (ii) the setting and the interaction style are based on metaphors of the orchestral Conductor and the theatrical stage.

The *story conductor*, i.e. the user of the system, is placed in a familiar setting to that of the orchestral conductor but in a virtual world modelled in a sense on the orchestra's stage. This stage serves as a visualisation of the context in which the conductor - user interacts.

This world has an open V shape, formed from three computer monitors. Multimedia objects feature on the two side monitors, and there is a "screen" in the centre of the stage i.e. on the central monitor. The functionality of the "screen" object is based on the context in which it operates, giving the user – conductor the sense of expectancy that all the "visual results" will appear on the screen.

The "instruments", i.e. *media objects*, manipulated by the story conductor represent a considerable range of media types, categorised as follows:

- 1. *A Sound Gallery*: a "jukebox" represents this sound gallery. Various sound objects, among them being music and environmental sounds, can be chosen.
- 2. *A Sound Effects Gallery*: a horn represents a series of sounds that can be used in combination with other sounds to emphasise points of the story.
- 3. *A Dialogue Gallery*: a picture of two people talking represents a series of short phrases that can be used in order to add appropriate voices to the story.
- 4. *A Film Gallery*: a camera represents a series of films (i.e. digitised video sequences).
- 5. *An Animation Gallery*: a cartoon character represents a set of characters that can be used in the story.
- 6. *A Photo Gallery*: A picture book represents a series of pictures that can be used in the story.
- 7. *Lighting Controls*: A light bulb represents the control of the lighting, for example to show the time of the day or night, or events such as sunrise, lights being switched on when entering a room, etc.
- 8. *Volume Controls*: A slide bar enables the user to change the volume of sound in any clip in which that sound is applicable, for example making a dialog be heard as a whisper or create a loud siren.

The above visual metaphors define the context of the conductor's world. The conductor's bimanual interaction is based on a vocabulary consisting of functional gestures, which are emblems by nature. Emblems are gestures that have standards of well formedness, a crucial language-like property that other types of gestures and pantomime lack [3]. Gestures are used in combination with eye tracking for media type selection. As implied by the conductor metaphor, the user interacts with the media objects in the same way as the conductor would interact with the musicians of an orchestra, i.e. establishing eye contact to initiate the interaction, and then using bimanual gestures to specify when and how the musicians will play. The orchestral conductor is silent throughout the interaction, but uses body language to convey information to the members of his orchestra. Based on this principle, silent interaction can be used to create multimedia stories. The user focuses on the media object to be

used and then uses gestures to select the particular clip. The story may be previewed during its creation and ultimately played in full. During the creation of the story, the conductor can choose the media that will be played, its ordering and other properties to be applied.

The goal of the user-conductor is therefore to create a story. The initial plan model is quite simple: select first media type, select first clip, select second media type select second clip, and so on. This process continues until the story is completed. However, the user-conductor may be more creative and may wish to adjust the presentation properties of the clips, the order in which they are played, and whether they are played in sequence or concurrently.

7. CONTROLLING MULTIMEDIA – THE VOLUME METAPHOR

Currently, designers must predict the types and level of information that users need unless their system can incorporate an advanced user-modelling system. The lack of a user-orientation in systems usually means that users expend considerable effort in adjusting the vast array retrieved information to suit their own requirements, both in terms of the format and level of detail.

Rogers and Scaife [4] have observed that students consistently admitted to ignoring text at the interface in favour of other media types such as diagrams and video material. Furthermore, when a variety of information is available (e.g. as is usually the case with web-based and multimedia systems), the user is left with the role of interpreting each individual representation and identifying any relationships between the different representations that may be present. Ainsworth [5] also notes that learners have difficulty translating between different representations and often fail to grasp important connections between different modes of representation.

The volume control under development by Zoë Swiderski is a mechanism that aims to address the issues outlined above by providing a basis for enabling multimedia objects to be controlled in an analogous manner to that of a volume control. A volume control (e.g. the Windows desktop volume control) is already a metaphor, which symbolises the notion of sound and provides a representation to enable users to increase and decrease its loudness. This metaphor could be effectively applied to controlling content levels of multimedia information.

Where volume levels can be 'turned up' or 'turned down', content level can also be 'turned up' by providing more detail (possibly by using additional distinct, but informationally related, media objects) or 'turned down' (using fewer media items and possibly less information). While addressing the issues of information overload by facilitating the filtering of information either by media type (images, text etc.) or detail (e.g. summary, full description, bullet points etc.), this mechanism has additional benefits. It could allow the expression of relationships between media types. Currently multimedia objects are interpreted as independent entities when it is more often the case that the collection of objects is being used to represent the same notion. Furthermore, if the combined information is examined, it can often present an alternative interpretation, which is lost when interpreting the objects individually.

8. CONCLUSION

This paper has described two systems that exploit multimodality, particularly gestural input. The story conductor features a more explicit use of gestural input, but in a sense, the "volume control" makes the control of detail a gestural activity. Furthermore, both systems exploit cross-modality and inter-medium reference, in that one form of input (gesture, adjusting a slide control, etc.) is converted into corresponding operations in an alternative media (order of shots in a video, level of detail in a text or diagram).

The two systems described reflect a "multi-modelling" approach to interaction whereby metaphors are used to represent relationships between input techniques, output presentation and the corresponding actions associated with their manipulation. This will enable users to exert full control over their digital environments.

The authors believe that the metaphors described in this paper can be seen as a step towards realising applications that are truly *multimedia* systems. They lay the foundations for exploiting not only media objects, but also the *relationships between them*. The systems described not only exploit analogous representations, but also analogous *processes (adjusting volume, conducting)* and *activities (controlling, directing)*.

Despite many claims being made for the power of metaphors in human-computer interfaces, there are few examples of the creative use of non-localised metaphors apart from the famous *desktop metaphor*. In this paper we have discussed the methodological background and the realisation of two systems, which transfer real-world interactions to novel metaphors thus bridging the gap between artificial environments and user interactions.

9. ACKNOWLEDGEMENTS

The authors are supported by the Distributed Multimedia Research Group. Zoë Swiderski's research is also supported by the Engineering and Physical Sciences Research Council (EPSRC) and British Telecom Laboratories.

10. REFERENCES

- Oviatt, S. (1999) "Ten Myths of Multimodal Interaction." Communications of the ACM 42(22), pp.74-81.
- [2] Eco, U. (1984) "Semiotics and the Philosophy of Language", Bloomington, U.S.A.: Indiana University Press.
- [3] McNeil, D. (1992) "Hand and Mind: What Gestures Reveal About Thought", University of Chicago Press, Chicago.
- [4] Rogers, Y. and Scaife, M. (1998) "How Can Interactive Multimedia Facilitate Learning?", Intelligence and Multimodality in Multimedia Interfaces: Research and Applications. J. Lee, AAAI. Press: Menlo Park, CA.
- [5] Ainsworth, S. (1999) "The Functions of Multiple Representations", Computers & Education 33, pp.131-152.

A semiotic communication model for interface design

Raffaella Scalisi PhD student in Communication Science and Design University of Turin Via Verdi 25 -10124 Torino - Italy 0039 349 8715200 scalisi@forcom.unito.it

ABSTRACT

This research wants to contribute to the creation of a semiotic framework for interface design. Using the Jakobson's communication model to analyse the HCI approach to interface development, we explain how some central factors of communication are not enough considered by designers.

Keywords

Communication science, human-computer interaction, interface design, semiotics.

1. INTRODUCTION

One of the basic assumptions of semiotics is that we cannot "not communicate" [24]. We always communicate even when we are not consciously sending a message. Semiotics underline that everything in the world communicates (from clouds to humans): *semiosis* is a pervasive phenomenon, but there is a difference between the clouds act of producing signs and the human act of communicating. In a simple *semiosis* the sender isn't well defined (clouds aren't conscious *adressers*), in the communication process the sender has a central role.

Our first question is what kind of communication is the computer's communication. It isn't a simple *semiosis*, but we must observe it's a special kind of communication, which, in some way, is more similar to advertising communication than human-human interaction. Using the Jakobson communication model, we observe how some communication factors and functions aren't well applied in the human-computer interface design and we suppose this fact may be a cause of some interface failure of effectiveness.

2. THE JAKOBSON MODEL AND THE HUMAN-COMPUTER INTERACTION

Ronald Jakobson proposed this model, taken from Weaver and Shannon cybernetic research, for his studies on poetic language [9]. Jakobson identified three basic factors (sender, receiver, message) and three other factors (context, medium, code). Each

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

factor has a communication function and Jakobson noticed that the meaning of a message can be produced by the weight of just one function, but any of them can be excluded.

We think that three communication functions\factors should be better studied in HCI: the *sender*, the *code* and the *context*. Before discussing these three factors, let's make some considerations on the others. We said we can't "not communicate", but there isn't any communication without a receiver and the meaning of a message is defined by the receiver's interpretation. Therefore the receiver has a central role. HCI studied deeply the receiver, in a non-semiotic approach, with the user-centred design theory, the cognitive studies on human understanding, etc.

User observation, user modelling, collaborative design, etc., are methods for knowing the user and developing his interface in an effective way. We think, however, that also semiotic studies on the *reader* can have an euristic function in HCI research field. We think of Umberto Eco *reader model*, his theories on *intentio auctoris*, *intentio operas* and *intentio lectoris* (the reader's interpretation based on his sign system, desires, emotions, beliefs, etc) [5]. The science of rethoric, literature and semiotics developed many methods and strategies "to guide" the reader interpretation. These strategies are ignored by computer science and most of cognitive studies, but we think they could be useful to improve human-computer interaction and they should be studied more deeply (metaphor generation is the only method that has been developed in this perspective [7]).

The medium is another central factor of the communication process. It's often said that the interface is the medium, but the medium is a very complex entity, especially in CMC [10] [18], that we should study separately.

2.1 The Sender

Semiotic HCI defines interfaces as messages sent by the designers to the users [4]. Can we really say the designer is the sender? Do users think the designer is the sender of the messages they receive during their computer sessions? If the answer is no, why we should use this model?

Do we have a real sender in this communication (if we think of a conscious sender)? Can we say the computer is the sender? Do users think about the computer as a real sender? The interface system is a world of signs in which we operate and manipulate, but sometimes computer also "speaks" to us. We think the user difficulty to understand the system resides in this second situation. Winograd and Flores [25] showed us how humans apply social rules in their relations with computers and how humans usually think of the computer as a conscious entity. The point is

that we should better clarify to the user, through the interface, the characteristics of this special sender (the computer) in order to avoid interaction problems derived from the misunderstanding of the sender's function. The aim of our research is to better understand how the user's conception of the sender influences the meaning interpretation.

2.2 The Code

The code is necessary for communication. Do we know codes from birth? No, we have to learn codes: from the meaning of clouds to the meaning of a sentence. Also the iconicity of the visual signs needs a learning phase to be understood. Think of the desktop icons: in semiotics this term describes a particular kind of sign which has the signifier similar to the content (see Peirce's theory on the three forms of sign representation: iconic, indexic, symbolic [16]), in interface design icons are similar to the object they represent, but we need a phase of learning to understand this codification. HCI aims to create systems we can use without learning, but we think this approach is too radical. We are able to learn, so why completely avoid this process?

Ease of use must be viewed as ease of learn. Creating systems with analogies to things we already know is a good way to reach this objective (desktop metaphor, infodomestics) [15], the creation of code conventions is a way to support understanding (the Web is adopting many conventions like "lens" for "search"). We have to continue working on codes, knowing that from a semiotic point of view the communication process is enabled by a previous knowledge of the code by the receiver. The science of rhetoric and linguistics can help HCI to create simple interface languages.

2.3 The context

The context has the power to change the meaning of a message (if you say "it's cold" and you are on the train and the window is open, you mean "close the window"; if you say the same sentence and you are in the mountains you just mean the weather is cold). The context is so important that Jakobson use this term to indicate the content. In interface design the context is the user's world, culture, job, computer knowledge, etc., but it's also the user's conception and understanding of the digital environment in which he operates. For this reason we should create coherent contexts in computer systems (the numerous studies on consistency in interfaces underline this need), but the problem is that the context change during, and in consequence of, a conversation\interaction. How can we build such dynamic contexts in a computer system?

The context has been studied in scenario-based design [12] and participatory design at the interface development level, but the concept of context should be studied more deeply. There are some very interesting approaches [11], [22], [25], [8] and, between them, we think Laurel's theories of user engagement and *interface as mimesis* are examples of a deeper insight. The field of computer game development should be considered too. [3].

3. CONCLUSIONS

HCI has focused his attention on the *medium* and on the *receiver*, but there are some factors in the communication process that are even more important: the *sender*, the *code* and the *context*. This paper refers to a research project that will study the effects of

these three factors on human-computer interaction, and puts in evidence some questions that HCI and semiotics should reflect on.

4. REFERENCES

- Andersen P.B., What semiotics can and cannot do for HCI, in CHI'2000 Workshop on Semiotic Approaches to User Interface Design.
- [2] Clarke A., Mitchell, Taking and making meaning: semiotics and new media, in proceedings of the First Workshop on Computational Semiotics for New Media, June 2000 University of Surrey, UK
- [3] Crawford C., Lessons from computer game design, in Laurel B., "The art of human computer interface design" Addison Wesley, 1990
- [4] De Souza C.S., The semiotic engineering of user interface languages, in International Journal of Man-Machine studies. No 39, 753-773, 1993
- [5] Eco U., Lector in fabula, Bompiani, Milano, 1979
- [6] Eco U., Trattato di semiotica generale, Bompiani, Milano, 1975
- [7] Erickson T.D, Working with interface metaphors, in Laurel B., "The art of human computer interface design" Addison Wesley, 1990
- [8] French T., Smith A., The contribution of semiotics to web-site design, in proceedings of the First Workshop on Computational Semiotics for New Media, June 2000 University of Surrey, UK
- [9] Jakobson R., Saggi di linguistica generale, 1966
- [10] Johnson S., Interface culture, HarperSanFrancisco, 1997
- [11] Laurel B., Interface as mimesis, in "User Centred system design: new perspectives on human-computer interaction" edited by D.A. Norman and S.W. Draper, Hillsdale, NJ Erlbaum Associates, 1986
- [12] Monteiro C.C., Barbosa S.D.J., de Souza C.S., The role of designer-generated scenarios in developing web applications: a case study, in Proceedings of the III Workshop on Human Factors in Computational Systems, IHC'2000, Gramado, RS. October, 2000
- [13] Nack F., About the influence of computer semiotics on communal intelligence, in proceedings of the First Workshop on Computational Semiotics for New Media June 2000 University of Surrey, UK
- [14] Nadin M., Interface design and evaluation semiotic implications, in Hartsin R., and Hix D., Advances in Human computer Interaction, Volume 2, 45-100. Norwood, NJ Ablex 1988
- [15] Norman D., The invisible computer, the MIT Press Cambridge MA 1998

- [16] Peirce C.S., Collected papers of Charles Sanders Peirce, Harward University Press, MA, 1931-1958
- [17] Prates O.P., de Souza C.S, Barbosa S., A method for evaluating the communicability of user interfaces, in ACM interactions, Jan-Feb 2000. Pp 31-38.
- [18] Prates O.P., de Souza C.S., Bicharra Garcia A.C., Semiotic framework for multi-user interface, in SIGCHI Bulletin Vol 29 N°2 1997
- [19] Prates, R.O.; Barbosa, S.D.J.; de Souza, C.S. (2000) A Case Study for Evaluating Interface Design through Communicability, in Proceedings of the ACM Designing Interactive Systems, DIS'2000, Brooklyn, NY. August, 2000
- [20] Reichman R., Communication paradigms for a Windows system, in "User-Centred system design: new perspectives on human-computer interaction"

edited by D.A. Norman and S.W. Draper, Hillsdale, NJ Erlbaum Associates, 1986

- [21] Saussure F. de, Cours de Linguistique Général, Paris, 1916
- [22] Sellen A., Nicol A., Building user-centred On-Line help in Laurel B., "The art of human computer interface design" Addison Wesley, 1990
- [23] Volli U., Manuale di semiotica, Laterza, 2000
- [24] Watzlawick P., Beavin J.H., Jackson D.D., Pragmatic of human communication, W.W. Norton & Co., New York, 1967
- [25] Winograd T., Flores F., Understanding Computers and cognition –a new foundation for design, Norwood, NJ Ablex, 1986

Life/Style OnLine[©] A Web-Based Methodology for Visually-Oriented Consumer Research

Michael R. Solomon Auburn University 308 Spidle Hall Auburn University, AL 36849 U.S.A. (01) 334 844-1316 solommr@auburn.edu

ABSTRACT

We describe $Life/StyleOnline^{\odot}$, a web-based interactive data collection technique that allows respondents to manipulate visual images of products as a means of expressing their tastes and preferences. This research tool is comprised of a browser-based software interface with an extensive database layer, which handles storage and retrieval of visual images.

KEYWORDS

Consumer research, online data collection, visual imagery

INTRODUCTION

The Web offers exciting new possibilities to communicate with consumers. Numerous commercial websites offer visitors engaging, interactive platforms, but the semiotic dimensions of presenting visuals online are still murky. In particular many social science researchers are not fully exploiting the capabilities of online media to probe deeply into consumers' motivations and preferences. Existing methods of data collection, largely focused on the written or spoken word, need to adapt to the enhanced possibilities for visual research offered by the Web. The use of visual stimuli by consumer researchers has largely been confined to small-sample qualitative studies that typically build upon methodological traditions developed in disciplines such as clinical psychology, visual sociology, aesthetics and anthropology. In the main, these techniques use visual material as part of a stimulus or response format and require "deep" interpretation as the analysis approach.

The *Life/Style OnLine*[©] technology was initially developed to explore the germination and dissemination of style trends among young fashion-forward female consumers and to contextualize their choices in fashion categories to other lifestyle choices. The project was funded by The National Textile Center, U.S.

First published at COSIGN-2001,

10-12 September 2001, CWI,

The Netherlands.

Basil G. Englis Berry College Campbell School of Business Mt Berry Station, GA 30149 U.S.A. (01) 706 290-2645 benglis@campbell.berry.edu

Department of Commerce. Although these web-based tools were developed in the context of a specific substantive domain, they have considerable general utility to other applications, both academic and commercial.

At the core of the project was the development of a web-based interactive data collection technique that allows respondents to manipulate visual images of products as a means of expressing their tastes and preferences. This research tool is comprised of a browser-based software interface with an extensive database layer, which handles storage and retrieval of visual images. One innovative feature of the software that generates the web pages is that it is a form of dynamic html (html scripting). Therefore, the large number of web pages that might be required by a specific research application are not individually created html files, but are instead interactively created online in response to the behavior of the respondent. The specific pages, then, do not actually "exist" until the time of application when the program creates them "on the fly" on the basis of respondent behavior, research design parameters, and information in the database (visual and verbal).

A major focus of our ongoing research program has been to learn how female fashion opinion leaders integrate information from mass-media lifestyle depictions as they form their own consumption preferences and communicate these choices to others. The basic feature of the product selection and assessment portion of the *Life/Style OnLine*[®] protocol is to have respondents assemble – on their computer screens – collages of product images they associate with images of selected people and lifestyle scenarios.

There are three visual layers in the current data collection paradigm: (1) sorting and selection of images of people in their daily lives (this involves the online sorting of images into discrete categories and the selection of category prototypes); (2) establishment of a social context in which product selection will occur; and (3) selection of an "ensemble" of products perceived to be ideally suited to each social context. This final step involves the selection of an "ideal" product from a larger set of category possibilities and the progressive assembling of a product ensemble. Both sorting and collage creation tools also pair image manipulation and selection processes with verbal response protocols (open- and close-ended formats). A demonstration version of the software is available at http://fafnir.berry.edu/ConsumersOnLine/index.html.