PROCEEDINGS

COSIGN 2003

3rd Conference on Computational Semiotics for Games and New Media

10 – 12 September 2003
University of Teesside, UK

ISBN

Copyright © 2003, University of Teesside

Printed in the UK

PROCEEDINGS

COSIGN 2003

Local Chair

Clive Fencott, University of Teesside, UK

Organizing Committee

Andy Clarke, Kinonet Consultancy, UK

Clive Fencott, University of Teesside, UK

Craig Lindley, Interactive Institute, Sweden

Grethe Mitchell, University of East London and Kinonet Consultancy, UK

Frank Nack, CWI, The Netherlands

Proceedings Editor

Grethe Mitchell, University of East London and Kinonet Consultancy, UK

Programme Committee

Elisabeth André, University of Augsburg, Germany

Wilton Azevedo, Mackenzie Presbyterian University, São Paulo, Brazil

Paul Brna, Northumbria University, UK

Kevin Brooks, Motorola Human Interface Labs, USA

Andrew Clarke, Kinonet Consultancy, UK

Brody Condon, Media artist, USA

Chitra Dorai, IBM, USA

Clive Fencott, University of Teesside, UK

Troels Degn Johansson, The IT-University of Copenhagen, Denmark

Werner Kriechbaum, IBM, Germany

James Lester, North Carolina University, USA

Craig Lindley, Interactive Institute, Sweden

Grethe Mitchell, University of East London and Kinonet Consultancy, UK

Frank Nack, CWI, Netherlands

Paolo Petta, OFAI, Austria

Mirko Petric, University of Split, Croatia

Doug Rosenberg, University of Wisconsin-Madison, USA

Ola Stockfelt, University of Goteborg/University of Skovde, Sweden

Robert Wechsler, Palindrome Inter-media Performance Group, Germany

New Media Artworks Committee

Andrew Clarke, Kinonet (UK)

Emmanuel Maa Berriet, La Graine/The Seed, France

David Daniels, Digital Poet, Berkeley, USA

Scott deLahunta, Writing Research Associates, Netherlands and Dartington College of Arts, UK

Fatima Lasay, University of the Philippines, Philippines

Luca Marchetti, Semiotician, France

Grethe Mitchell, University of East London and Kinonet, UK

Timothy Portlock, University of Illinois, USA

Doug Rosenberg, University of Wisconsin-Madison, USA

Table of Contents

Preface

Welcome to COSIGN 2003, the third COSIGN conference. Once again, we have brought together artists, researchers and academics with the intention of producing the unique mix of people and ideas that gives COSIGN its unique atmosphere.

COSIGN 2003 is the biggest COSIGN event so far, spanning for the first time three full days. We are delighted to have Ernest Adams, Marc Cavazza and Lev Manovich as keynote speakers, and in addition to the stimulating papers, demonstrations and presentations, we have scheduled plenty of time for both formal and informal discussion as this has been such a key component of the previous two conferences.

COSIGN started from the notion that the way in which we take and make meaning with computers is becoming an increasingly important issue and that semiotics could provide a critical component of this study. The success of the COSIGN events - and strength of the papers and artworks submitted have - in many ways, vindicated our faith in this original idea. We hope that at the end of this conference, everyone will be able to go away with a better understanding of the field of computational semiotics, and with new ideas and techniques that they can feed back into their own work.

Producing a conference requires the effort of many individuals working together towards a common goal. We would therefore like to thank all of the people who have volunteered their time and talent in support of COSIGN 2003. We would like to thank the University of Teesside for its help with resources, rooms, technicians, etc. Particular mention should go to Carol Cook of the University's Research and Enterprise Office for her unstinting support and enthusiasm, and to Janice Webster, director of the VR Centre, for both support and for facilitating delegate visits to the centre.

Within the School of Computing, thanks must go to Jim Yip, Director of the school, and Derek Simpson, Assistant Director, for their continued support and practical advice. Invaluable assistance was given by the team within the Administration Office of the School of Computing and Mathematics - in particular Karen Shaw (who handled the registration process) and Caroline Walsh, Emma Howden and Lynne Dawson (who took care of many other arrangements so necessary for the smooth running of a conference such as COSIGN).

We are grateful to Elaine Bradwell, Service Support Coordinator, and the Repographics Department for their help and assistance in printing the proceedings. We would also like to thank the staff of the Centre for Enterprise at the University of Teesside for their help both before and during the event, and Cleveland College of Art and Design for the loan of temporary display panels. If we have forgotten anyone, we hope that they will forgive us.

A conference also requires an operating budget, and we gratefully acknowledge the support provided to us by the sponsors of COSIGN 2003 - it is their generosity that has allowed us to keep the registration fees at an appropriate level. In addition to University of Teesside, who provided financial and organisational support for the conference, we would also like to thank Digital City not only for their financial support, but also for their general help and advice.

Final thanks must, of course, go to you, the participants of the conference. Without the energy that you bring to the conference, and your willingness to share ideas, none of this would be possible.

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

Ernest Adams

Transmitting Meaning In Interactive Contexts

Computer games are not inherently linguistic, inherently assertive, or inherently narrative. It is possible to construct entirely abstract ludic experiences that are devoid of any kind of symbol whatsoever. It is also possible to construct games using traditional narrative techniques (though whether they constitute a "literary" narrative, whatever that means, remains the subject of debate), essentially leading the player through a story which he may choose to interpret using the conventional methods of literary criticism. However, it is also possible to construct yet a third type of game, a non-narrative game which nevertheless conveys assertions to the player purely through his interactions with the game's symbolic elements. That is to say, the game makes no overt assertions, but is constructed in such a fashion that the player's interaction with the game produces the impression of an assertion in the player's mind. This is, I believe, the great challenge to developers and semioticians alike: understanding the transmission of meaning in interactive contexts. It is this third type of game that I intend to examine in this lecture, drawing examples from the commercial game industry, including such works as Sim City, Balance of Power, Catch the Sperm, and America's Army. I do not propose to offer answers, but I hope, perhaps, to help frame the questions.

Biography

Ernest Adams is a 14-year veteran of the game industry, now working as an independent game design consultant, writer, and teacher. He is the author of two books, "Andrew Rollings and Ernest Adams on Game Design," with Andrew Rollings; and "Break Into the Game Industry: How to Get a Job Making Video Games." Ernest was most recently employed as a lead designer at Bullfrog Productions on the Dungeon Keeper series, and for several years before that he was the audio/video producer on the Madden NFL Football product line. He has developed on-line, computer, and console games for everything from the IBM 360 mainframe to the Playstation 2. He was also the founder of the International Game Developers' Association. His professional web site is at http://www.designersnotebook.com

Marc Cavazza

Alternative Reality: Can Virtual Reality Art Inspire VR Research?

Early work in Virtual Reality was characterised by a quite radical stance, such as Tim Leary's "Psychedelic" vision of VR. While this vision has tainted the popular culture image of VR, most scientific work has been dedicated to the pursuit of realism or, short of achieving sufficient accuracy, the illusion of presence. The greatest number of virtual worlds produced to date is a product of the many genres of video games: they boast sophisticated Physics engines, which, together with ever improving real-time graphics, underpin the quest for realism.

What is left, then, of VR as an endeavour to question our everyday experience; as a means to explore alternative realities?

Several VR Artists, such as Char Davies, Toshio Iwai or Maurice Benayoun, have explored in their work the concepts of time, space, as well as the laws of Physics. For instance, the famous animation series "The Quarxs", ideated by Benayoun, featured virtual creatures that would defy the established laws of Physics and, in some cases, even transform the validity of these laws in the environment they populate.

This work inspired us into exploring the kind of techniques that could support the definition of alternative physical behaviour in virtual worlds. Explicitly redefining alternative laws with new equations to be solved analytically is not a practical option, especially if alternative behaviours are to coexist with traditional ones. What is required is a high-level formalism that could represent physical processes themselves: alternative laws of Physics could then be represented in a principled fashion using a similar model. Such formalisms do exist: they have been developed in Artificial Intelligence (AI) within a field known as qualitative physics, whose objective is to describe physical processes in a symbolic fashion. Qualitative Physics is based on discretised descriptions and as such is compatible with the event-based architectures used in modern game engines.

Another important concept for our apprehension of reality is causality. In addition, the cognitive attribution of causality is an important element of human experience, hence a privileged target for experiments in alternative reality. As causality is attributed mostly from the co-occurrence of events, one possible research direction to manipulate causality in virtual worlds is to intercept low-level events and transform these prior to their re-activation. The process by which these events can be transformed and re-ordered bears analogies with some AI planning techniques. This has the potential to open new research directions, departing from traditional work on causality.

Biography

Marc Cavazza is Professor of Intelligent Virtual Environments at the University of Teesside. His current research explores AI-based simulation techniques in virtual environments to implement alternative laws of physics and causality. Over the past years, he has explored the use of most AI techniques in virtual environments: search, planning, natural language processing and qualitative physics. He has authored over 120 publications in various fields of AI and VR. He is currently co-ordinator of the ALTERNE project, an IST project developing technologies for VR Art.

Lev Manovich

"Metadating" the Image

Metadata is the data about data: keywords assigned to an image in a media database, a number of words in a text file, the type of codec used to compress an audio file. Metadata is what allows computers to "see" and retrieve data, move it from place to place, compress it and expand it, connect data with other data, and so on. For instance, metadata identifies a set of numbers stored in a file as as an image rather than a spreadsheet to the software. More generally, metadata determines the semantics of data within the computer.

Modern age is characterized by the ongoing struggle between the visual data, i.e. images, and their creators and masters - the humans. The later want to control images: make new images which would precisely communicate the intended meanings and effects; yield the exact meanings contained in all the images already created by human cultures; and, more recently, automate these and all over possible image operations by using computers. The former can be said to "resist" all these attempts. This struggle have intensified and became more important in a computer age - more important because the ease with which computers copy, modify, and transmit images allows humans to daily multiply the number of images available.

"Metadating the image" is not, however, only the economic and industrial problem to be solved - it is also a new paradigm to "interface reality" and the human experience in new ways. This is already demonstrated by a number of successful art projects that focus on new ways to describe, organize and access large numbers of visual records. Importantly, these projects propose not only new interfaces but also new types of images, or, more generally, "records" of human individual and collective experience: panoramic video recordings (ongoing projects by Jeffrey Shaw and Luc Courchesne); film/video recordings embedded within virtual space (Sauter, Invisible Shape of Things Past; Fujihata, Field-Work@Alsace); photographs of people/objects organized into networks/maps based on their semantic similarity (Legrady, Pockets Full of Memories; Walitzky, Focus).

In summary, in terms of its creative and "generative" potential, "metadating the image" paradigm means following four related directions: (1) inventing new systems of image description and categorization; (2) inventing new interfaces to image collections; (3) inventing new kinds of images which go beyond such familiar types as "a still photograph" or a "digital video"; (4) approaching the new "super-human" scale of visual data available (images on the Web, web cam recordings, etc.) not as a problem but as a creative opportunity.

In short: new structure - new interface - new image - new scale.

Biography

Lev Manovich is an Associate Professor in the Visual Arts Department, University of California, San Diego where he teaches courses in new media art and theory. He has been working with computer media as an artist, computer animator, designer, and programmer since 1984. His art projects include: little movies, the first digital film project designed for the Web (1994); Freud-Lissitzky Navigator, a conceptual software for navigating twentieth century history; Anna and Andy, a Web adaptation of Anna Karenina; and a digital film project Soft Cinema commissioned for ZKM exhibition Future Cinema (2002-2003). He is the author of The Language of New Media (The MIT Press, 2001), Tekstura: Russian Essays on Visual Culture (Chicago University Press, 1993) and over 50 articles. Currently he is working on a new book, Info-aesthetics. His website is http://www.manovich.net

Computational Semiotics - A Growing Interdisciplinary Field

Grethe Mitchell, University of East London
and Kinonet Consultancy
g.r.mitchell@uel.ac.uk grethe@kinonet.com

Andy Clarke, Kinonet Consultancy, UK
andy@kinonet.com

Frank Nack, CWI, Netherlands
Frank.Nack@cwi.nl

Clive Fencott, University of Teesside, UK
P.C.Fencott@tees.ac.uk

Craig Lindley, Interactive Institute, Sweden
Craig.Lindley@interactiveinstitute.se

The papers and presentations of the third COSIGN conference on Computational Semiotics in Games and New Media illustrates the growing interest in issues dealing with the creation, negotiation and application of meaning in digital interactive media – and the concomitant growing development of semiotic theory dealing with computational communication.

As digital media itself grows more complex and its different ramifications start to spawn new modes and new ideas, so too does the research in how we both create and decode meaning in digital interactive media, and in this area semiotics (the study of signs) has once again become relevant. After languishing for some years on the periphery of postmodernist thought, semiotics has perhaps found its ideal medium in the multiple-media, computer-mediated responsive technology whose existence and performance relies upon layers of interacting code and codification.

At the heart of the semiotic project lies the conceptualisation of sign systems as social constructs, that is to say, as products of the society and environments in which the sign systems operate. Thus semiotics could initially be construed as an offshoot of sociological or anthropological theory, and therefore, from the point of view of science, possibly not suited to the supposedly neutral world of computation, although very useful in social and cultural research. However, within the 'broad church' of the semiotic project there is also a preoccupation with structure and with defining the structural elements of sign systems that hold true for a given language, whilst simultaneously accepting the idea of social construction. This structural yet social view of sign systems was fist developed by Levi-Strauss, and then elaborated by Barthes, Eco. Pierce and others from the foundation of linguistics. Given that any sign system has to operate within a given social environment, it would seem illogical not to accept the social dimension of communication, yet this acceptance need not interfere with the elaboration of the structural dimension of sign systems. It therefore becomes possible, not to mention useful, for semiotic theory to be applied to computational communication, both as a procedural mechanism for developing communication systems and as a model for the analysis of the effects of sign systems upon users and their environment.

The usefulness of semiotics in providing a theoretical framework in which to discuss, analyse and implement meaning within the field of Artificial Intelligence (AI) is discussed in the paper by the artist and AI researcher Michael Mateas. In 'Expressive AI: A Semiotic Analysis of Machinic Affordances', the author makes the case for the usefulness of structuralist semiotics as a model for exploring the relationship between interpretative and authorial affordances described in his formulation of the relationship between the negotiation of meaning and the internal structure of an AI system, the practice of which he names 'Expressive AI'. Mateas argues that Semiotics can assist with what is often described as the 'knowledge representation problem': i.e. the task of defining structures and processes that are amenable to computation whilst at the same time being meaningful to the user. Using a semiotic model to define both sides of the 'knowledge representation problem' as instances of sign systems, Mateas sets about analysing and unpacking the relationship between the 'code machine' (sign system of code) and the 'rhetorical machine' (the sign system used to talk about the code). Whilst acknowledging that it has its limitations (which could in time be addressed), Mateas nevertheless makes a case for the usefulness of structural semiotics in providing both a common ground between the two sides of the 'knowledge representation problem' as well as a framework for more detailed analyses of the relationship between the authorial and interpretative affordances of Expressive AI, and hence between the code system and audience interpretation.

A structuralist approach is also used by Szilas in determining a procedural and atemporal model of dramatic narrative in order to create a new form of digital interactive art where the user participates in the narrative (designed by the author) by taking the role of one of the characters. Drawing on the structuralist

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

writings of Eco, Bremond, Barthes and Todorov amongst others, to create a narrative logic combined with a narrative sequencer and a user model of narrative perception; Szilas presents a demonstration of his Interactive Drama system: 'IDtension'. As there are no predefined narrative threads in IDtension, the story unfolds as a result of interaction and thus reveals itself only at the end. For the author, (and as discussed in his paper presented at COSIGN 2002) structuralism must play a central role in the system as it provides the sole methodology for determining narrative structures designed to be simulated rather than reassembled.

In 'Contact Expressions for Touching Technologies', the authors McGee and Harup acknowledge the usefulness of semiotic theory, whilst also decrying the fact that up to now there has been surprisingly little investigation into the semiotics of touch and 'contact expression'. They go on to make the case for the development of this area of study and application, in the context of computational interfaces, by discussing their research into HCI and interface design for tactile technologies and outlining their experimentation with a 'contact cushion' used to explore some of the potential for contact expressive devices. With the goal of contributing to the idea of 'syntonic design' (as described by Papert) for empowering computational devices, applications and services by outlining an initial design taxonomy and contact-expressive design Pattern (after the work of Christopher Alexander), the authors instigated a cycle of implementation, study and revision (using the Pill O' Mate responsive cushion device) – where the implementations are based on cognitive insights from the intersection of developmental ('constructivist') psychology (as in the work of Piaget) and semiotics. The outcome of the research so far suggests that it would be fruitful to further the understanding of the cognitive semiosis of contact expressions in order to benefit the development of empowering contact-expressive devices.

The application of semiotics to HCI and interface design is also the concern of O'Neill and Benyon. In their paper, 'A Semiotic Approach to Investigating Presence', the authors suggest that Semiotics can become a new foundation for interactive systems design (if brought up to date), as it can offer a highly appropriate level of discourse in this area. The authors argue that, differing from the traditional focus of HCI on efficiency of tasks and details of buttons and menus, they wish to investigate the interpretations, meanings and significances experienced by the user of interactive systems. Drawing, amongst others, on the writing of Umberto Eco and Peter Bogh Andersen, on Paul Dourish's theories of 'embodied interaction' and the idea of 'umwelt' (proposed by Jacob von Uexkull) to explain their position, they go on to describe how semiotics can help interpret the findings of their investigation (as part of the BENOGO project) into 'presence' in virtual and real environments. Comparing the reactions of users to both a physical ('real') and a virtual environment of the same subject (a botanical garden), the authors were able to determine instances of connotation within the virtual environment (alongside the expected denotative interpretation). For the authors, semiotic theory provides a framework in which to identify and discuss the mechanism of connotation in virtual and real environments, as well as determining the temporal and spatial relationship between signs, the links between objects within the environment and the signification process that they engender. Semiotic analysis reaches beyond denotation, to provide insight into the connotative area of signification, and opens up HCI and interactive systems design to these considerations.

We have seen that the semiotic theories provide useful models and frameworks for the construction and design of computer systems and the concerns of HCI. As discussed above, a further dimension of semiotics lies with its application to analyses of social behaviour and discourse. Acknowledging that language (or any sign system) is not neutral, but rather a product of both social convention and individual experience within society, semiotic analyses of communication media, cultural practices and products provides us with useful insights into the relationships between the individual, technology and the cultural or media output of a given society, as evidenced by the writers such as Roland Barthes and more recently, Kress and Van Leuwen. As computational media is simultaneously technology, object of consumption, cultural practice and artifact, it becomes inevitable that semiotics would also be gainfully applied to the social and cultural aspects of computer use and interaction.

In their paper 'Signs From a Strange Planet: Role Play and Social Performance in Anarchy Online' Burn and Carr draw upon the theories of social semiotics elaborated by Kress and Van Leuwen to investigate and analyse the numerous modalities at work in the online role playing game Anarchy Online. Social semiotics conceptualises signs as arising from and constructing social discourse. With signs thus discursive and motivated (rather than arbitrary), the authors' investigations centre upon the motivations at play within the on-line role-playing game (RPG) 'Anarchy Online'. In doing so the authors determine the existence of representational, ludic and communal motivations and outline the semiotic activity related to the category of motivation, noting that some activities straddle more than one category. Semiotics also provide the authors with a tool for describing the manner in which some meaning generation is completely player-produced (e.g. the chat mode content which itself follows the linguistic conventions of *txt*) whilst other semiotic activity is produced by a combination of player, game design and other players. Categorising the semiotic systems within the game allows these to be analysed and discussed, as well as determining which aspects of semiotic activity are part of the game and which are infiltrated by concerns outside of the RPG. The semiotic analysis of the multi-modal interaction of the game also assists with determining whether these either reinforce or work against the locale of the game and whether this matters.

As alluded to above, computational communication demands an awareness of the multiplicity of semantic codes and the layering and interaction of codes from the lowest level upwards to the graphic, physical, spatial and aural languages of the interfaces and interactions. No wonder then that numerous practitioners and researchers are pre-occupied with the issues of language and comprehension in computational communication. Of these, successful speech recognition and natural language processing still presents a 'holy grail' quest – in that finding the solution is highly desirable yet immensely complex and difficult. Nevertheless, the field of semiotics provides useful models for analysis and for the construction of devices to explore and solve issues of language processing.

For example, in their paper 'A Game Pidgin Language for Speech Recognition in Computer Games' the authors Rudra, Kavakli and Bossomaier present the case for the development of a language designed to tackle the problem of speech in human-computer interaction: Game Pidgin Language (GPL). Positing the use of a pidgin language structure the authors present it as a solution for the simplification of natural language so that it may be used in natural language processing, with specific reference

to its use in speech for computer games. Whilst conditions of pidgin language, such as limited vocabulary and simple grammatical structures assist in minimising the difference between the Entropy and Perfect Information content of the language (these same conditions also being appropriate for the use of such a language in computer games), the authors also emphasise the need to focus on the semiotic qualities of the pidgin language – explaining how, in order to create a successful computer pidgin language, each word needs to carry a significant semiotic weight so as to achieve the goal of a good match of word with utterance, or in semiotic terms, of *langue* with *parole*.

As evidenced in previous years, the COSIGN series of conferences are unique in the manner in which they bring together practitioners and researchers from the realms of science and the arts in an interdisciplinary format, replicating the manner in which successful computational communication also necessitates an engagement with scientific principles and with artistic investigation. The study of, and production of meaning in, computational technology demands awareness of factors pertinent to science and to the arts and it therefore seems logical that both scientists and artists are consulted on these matters.

Thus the relationship between language and computational communication is also the theme of new media artists Horowitz and Deck who, in their work, reflect an engagement with the semiotic project of decoding media to tease out underlying philosophical and political concerns – which in these works are pertinent to the use and status of language in computational media.

As seen above in the work by Rudra et al, matching utterance with word is a prime goal of speech processing. Utterance and the relation between the spoken word and its computational representation is also the concern of Risa Horowitz's piece: 'Melitzah' (meaning utterance in Hebrew). The work uses computer technology for storing and displaying the graphical representation – in waveform - of every word in the Canadian Oxford Dictionary. As such it is a visual translation of the English language, a database of sounds and waveforms representing words; a visual representation of the auditory. It is also, in the audio recording, a representation of the vocal inflexions of the artist. In displaying the waveform pattern for spoken and written words, Melitzah compares the physical manifestation of utterance (waveforms) with the arbitrariness of its linguistic signifiers (words) – underlining the extent to which the production of meaning and its understanding is a construct. Melitzah also questions the 'empowering' and 'open-ended' status often attributed to interactive art works, as it is clear that the work is constrained by its vocabulary – large though this is – and the authorial intentionality of the artist. The artist maintains that interactive art is no less controlling of the user than other well executed art forms. Nevertheless, Melitzah does provide space for the user to participate – creatively – in the work and thereby 'acquire a sense of authority and authorship over their interaction with the work'.

Issues of control and authorship in relation to language and computation are also addressed in Andy Deck's artwork: Lexicon. Deck describes Lexicon as addressing 'the relationships between verbal and software languages, and between free expression and coercion'. As a form of software that needs to be 'discovered rather than immediately understood', Lexicon mirrors the process of language learning, just as it mirrors the process of signification when offering a series of encounters with 'programmatic images' as the application translates words into imagery. Lexicon also questions aspects of human computer interaction, using these to frame wider questions about our control and participation in the writing and signification of computer language and how software generally flows downwards from experts to consumers. Whereas spoken language is the prerogative of most, computer languages require specialised initiation. Deck makes the point that Lexicon is designed to expand its vocabulary but that this does not avoid the issue of expert knowledge – and Lexicon illustrates this by allowing development of its vocabulary on condition that the developer either has knowledge of the Java programming language or is sufficiently ambitious. In pointing to the 'linguistic character of software' and hence the possibility of computer languages to be extensible and participatory sign systems, Lexicon also provides contemplation of the importance of the open source movement in facilitating freedom of expression.

The fruitful collaboration between science and art in the field of semantic implementation is further displayed in the paper 'A Recombinant Information Space' which outlines the creation of innovative browsing software 'CollageMachine'. The authors, Kerne and Sundaram, draw upon the practices of the Dadaists, Duchamp and Max Ernst to illustrate how methods of collage and recombining, that is to say, decoding and re-coding existing texts, can be a useful model for the development of an innovative interface for the production of meta-documents which serve to improve both browsing activity and data classification. In addition to the artistic practice outlined above, Kerne and Sundaram also draw upon 'hard science' image processing techniques to create visual layering that further assists with sifting and prioritising information - thus providing an example of how the construction of CollageMachine is dependant on the interaction between artistic practice and scientific formulation.

In a further, yet alternatively motivated, collaboration between art and science, the artist Andrea Polli uses computer technology and scientific theories of perception and vision to inform her interactive installation 'The Fly's Eye', drawing particularly upon the biological structure and function of the fly's eye to deconstruct live video images of both the audience and the space in which the audience is located. At the same time as being observed, the audience also has some control over the outcome, depending on their movement. In the real-time computer manipulated deconstruction and analysis of video images 'The Fly's Eye' challenges the accepted codes of representation. In so doing it opens up possibilities of new semantic arrangements and hence a 'new reading of physical information'.

Also at the intersection between science and art, the artist Ivo Flammer:, in '850 hPa 01 MEZ', presents the relationship between the so-called objective, neutral and precisely defined language of science and the multi-dimensional, emotional and personal language of contemporary art. He does this by combining the two languages in an interactive installation based upon meteorological charts, designed to raise the viewers' awareness of how their own decoding system, based upon their 'umwelt' or 'personal encyclopedia', subjectifies their observations.

Returning to the concerns of HCI and their implementation, the use of a visual language in this respct is also the concern of the researcher David Bihanic. His paper: 'A Complete System of Tridimensional Graphical Representation of Information:

"Crystal Hy-Map™"' presents work done on implementing a new visualisation for the organisation of metadata that does away with the now traditional WIMP (windows, icons, menus and prompts) model and introduces a 'tridimensional graphical representation' with which to explore information. Crystal Hy-Map™ uses a hypermedia structure in which information is indexed by geo-semantic analysis, positioning the information in space according to its semantic relationship to other units of information. The semantic distribution of information allows for the constitution of a 'semantic network' where browsing and specific searches are assisted by relationships of meaning.

The relationships of meaning are also a theme in the work of artist Ennio Bertrand who, in his project 'Sucker Today' combines stills of TV news footage and disconnected sound fragments with user interaction to create installations that query the manner in which news events are initially processed and mediated by the sender and then perceived and newly elaborated by the receiver. In 'Sucker Today' the user journeys through a virtual labyrinth composed of media stills of the demonstrations at the G8 Summit, held in Geneva in 2001 and which culminated in the death of Carlo Guiliani, a young man shot by a policeman. The combination of images and sound fragments - activated as the user passes across them - in a labyrinthine form, encourages the user to construct a personal narrative thus questioning the relationship of meaning between media production and consumption.

As we discussed back in 2001, computers are meta-data machines and computational media is now a meta-language communication mechanism, drawing upon a multiplicity of codes to express meaning and affect emotion. Computational communication brings into play a multiplicity of interacting signification: written and spoken language, visual and aural language, static and moving language, body language. Added to this are the communication codes derived from film, drama and choreographic techniques: set design, montage, performance in dance and theatre, camera angles and positions, movement through time and juxtaposition of signs, images and sounds. Also, graphic design and typography, sound sampling and mixing – all of this contributing a host of interrelated signs playing off and with each other to build meaning. Not forgetting the signification of interactivity – the meaning of response and counter response – a feedback loop in real time between user and used, between soft machine and software. The multiple codes used in the semiosis of the meta-medium of computer communication are also thematic concerns pertinent both to the conference and to the wider boundaries of computational semiotics.

This is reflected in the paper 'Symbolic Activities in Virtual Spaces' by Schäfer et al, the main concern of which is the application of structures of narrative and drama to computational communication and storytelling at the service of asynchronous collaborative working. In their paper, the authors outline an approach combining concepts of 'symbolic acting' and virtual storytelling to support the process of asynchronous collaborative working using virtual environments. The paper describes 'DocuDrama'; an application which supports collaborative working using dynamically generated storytelling strategies. DocuDrama uses three modi to present project activities and interaction between team members, with each approach using a symbolic language of its own. In order to facilitate easy comprehension and intuitive understanding the symbolic languages are developed drawing upon signifying systems and languages of pre-existing media, e.g. cinema, theatre, storytelling, body language – as well as using approaches from computer games such as The Sims.

From the perspective of a hybrid system combining theatrical performance with computational communication and drawing upon the work of Oskar Schlemmer and his idea of the body as 'code' – as abstract signifier rather than as representation of human emotion. expression or psychology, the author Wolf Ka presents a paper: 'From Text to Interface: Theatre and Digital Media', describing how the theatre company: res publica, in their performance 'Enjeux', use interactive computational technology to query the construction (and deconstruction) of theatrical representation, the formulation and arrangement of space, and the relationship between the stage and the audience.

By means of an 'interactive scenic device' consisting of an interface projected onto the stage and an input device in the form of a trackball passed around the audience, the performance becomes dependant on the input from the audience – who then become a constituent part of the work, integrated into both form and content. In addition the audience are also confronted with their sequence of choices when these appear as titles in a 'meta-text layer' projected onstage.

The projection queries the relationship between two different sign-systems: the graphical (and textual) interface and the given performance. The interface refers to graphic, linguistic and iconographic signs, whilst the performance refers to the expressive abilities and codes of the human body. Through computational communication the performance links these two systems in order to allow the audience to actively participate in the performance – 'to intervene in the process of representation'. In doing so the piece contributes to the enquiry of how technologies modify cultural practice in the context of theatrical performance, and it also serves to demonstrate and document that decoding the 'digital paradigm' is not necessarily a solitary activity in front of an individual screen, but can also be a collective experience.

We know that spaces are semiotic locales, and it is often difficult to change the inbuilt meaning of a virtual space – just as this is also difficult in the real world. In her paper 'Messages For a First Person Perspective', Maia Engeli discusses an exploration of the creation of alternative messages within the first person shooter game 'Unreal Tournament' – the results of two workshops with Interactive Media Students in Switzerland and Architecture students in The Netherlands. Engeli points out that FPS games allow the creation of very powerful 'messages' that are communicated by means of an expressiveness that is the combination of visual, aural and interactive elements in a 'dynamic digital space'. Emphasis in the workshops was given to exploring the expressive power provided by FPS and the results were exhibited for two weeks. Discussing user reaction, Engeli outlines how UT players found it difficult to change their readings of the environments from FPS mode to the more experiential mode required to partake of the message and what would be required in order to bridge the gap. Engeli goes on to point out that the semantics of VR is currently still rooted in physical reality and that this constricts both the 'reading' and 'writing' of new 'messages'.

The multi-modal constituents of computer communication and the alternative realities of VR are also to be found in other practices and locations. The club scene spawned in the eighties can be seen to have a symbiotic relationship with computer generated communication, both of which became culturally embedded during the same time slot. 'Synaesthetic Performance

In the Club Scene' by Annet Deckker, makes the point that both VR and the DJ/VJ performance are attempts to take an audience (user) through Alice's looking glass into another world. As with the multiple modalities of computer media, DJ/VJ performances are synaesthetic events combining a number of signifying practices (sound, image, special effects, smell etc) to create an alternative reality. In her paper, Deckker provides an account of the development of VJ and outlines the historical background to the DJ/VJ event, linking it with performance art and with earlier examples of synaesthetic performance such as Wagner's Gesamtkunstwerk (a fusion of the arts) and the even earlier 'diorama' performances of Daguerre and the panorama painters of the 18th century whose goal was 'an active experience of space'. With similarities to interactive performance (see the paper by Wolf Ka) computer mediated interaction changed the audience from viewers into users, influencing as well as participating in the event. Live streaming internet connections link VJs to their public across the world – the computer modalities pulverizing distance and time, creating instantaneous signification for global clubbers.

Finally, perhaps because of our familiarity with both its ubiquitous nature and all pervading presence - and the language we use to describe the disembodied nature of the wired universe, (as opposed to VR), we are able to view the net in terms other than physically rooted in the Newtonian space referred to by Engeli. A new reading of both the cyborg entity and the net is provided by transmedia artist Christina McPhee, where the digital paradigm of the net – seen as 'inscape' – becomes both the object, site and subject of poetic speculation. In her paper, 'Aphasia + Parrhesia: Code and Speech in the Neural Topologies of the Net', McPhee uses the allegory of the cyborg as neural entity to delineate and comment upon the interacting webs of communication and code that constitute cyberspace. The author describes the net as a cyborg brain, exhibiting the effect of neural trauma in her disintegrating entropic memory combined with recursive persevering aphasia - where the discontinuity of language is interspersed with flashes of meaning and the cyborg nevertheless still attempts to 'reflect, react and remain autonomous', despite 'being at the service of the code and the writers of the code.'

The semiotic project has often been criticized as a wanton act of destruction - unraveling or unpicking a work so much that any meaning is lost. But the papers presented at COSIGN2003 highlight the many strengths of semiotics when applied to computer-based and computer-mediated works. In spite of their diverse approaches and subject matter, these papers – and those from the previous conference – are forming a coherent body of work, and it is becoming clear that computational semiotics has an important role to play: on one hand as a tool to create meaning with computers, and on the other as a technique with which to analyse meaning in computational media.

Yet the diversity of the papers presented show that computational semiotics is not just a set of rules that can be learned by rote and applied in any circumstance. It is something more elusive than that: a way of thinking, rather than a formula. Through the COSIGN conferences, the field of computational semiotics is being defined and shaped, building up a body of knowledge that will help to continue this process of discovery.

SuckerToday

Ennio Bertrand
Via Soldati 12
20154 Milan Italy
+39 - 02 349 5621
bertrandart@iol.it

ABSTRACT
This paper by Ennio Bertrand presents some of his interactive installations dealing with news media. It focuses on one of his latest works: *SuckerToday* (2002). This installation is dedicated to the clashes between demonstrators, black blocks and policemen which took place in Genoa (Italy) during the G8 summit in July 2001. It invites the user to explore media stills captured from news programs (organised in a labyrinth and accompanied by unrelated sounds and voices), through a classical videogame setting with joystick and screen.

Keywords
Interactive installation, news media, art

1. INTRODUCTION
A large part of my work focuses on news media, tv news especially. I'm very much interested in the way news are processed, mediated, and then perceived, memorized, newly elaborated by the audience.

The same is true for the interactive installation *SuckerToday* I'll present and discuss here. This work is part of a long lasting research I've been doing during the last 10-15 years, which I'll briefly sum up.

2. INTERACTIVE INSTALLATIONS PLAYING WITH NEWS
For many years (starting approximately in 1990) I've been capturing stills from tv news programs and documentaries. I am used to print them, usually covering unimportant parts of the images with a golden or silver leave (somehow like the golden medieval paintings) in order to put in front those details which I find more interesting. I often create sequences of up to 12 images, usually summarising un unnoticed little story happening during (and covered by) some bigger event: e.g., the men carrying Lady D's coffin, each of them looking directly inside the camera as soon as he enters the church where the funeral is to be celebrated. Through this time-freezing process I want to remove the images (and the events) from the tv flow, with its flattening effect, and to give them their dramatic consistency back.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

Figure 1: Lady # 2, 1998

I've started thinking of doing some interactive installations in 1992. For example *Lux sonet in tenebris*, that was built out of 144 little self-made loudspeakers, each with a built in light sensor. As soon as the sensor is covered by a shadow, the loudspeaker emits a little squeaking sound: so that a person can produce sounds through moving in front of the loudspeakers. The installation is very playful and easy to interact with. Another installation, *Smell Link* (1996), was meant to play with people's misunderstanding of what exactly we send through the web when we connect to the internet. It allowed people in Torino (Italy), Cuenca (Spain), Aix-en-Provence (France), Trieste (Italy), to exchange smelling portraits (with a perfume chosen by each user) to one another.

The very first interactive installation I did that uses sound and video images was *Memory of the surface* (1992-1998). *Memory of the surface* recalls the shadow which some person's body left on a wall in Hiroshima as a consequence of the H-bomb explosion, as documented by a famous photograph. A video-camera monitors the silhouettes of those who stand in front of a blank light cloth screen surrounded by debris. There is a crescendo of a gloomy sound, recalling the American bombers, culminating with a flash and a sudden silence. In this very moment those standing are photographed and their silhouettes reproduced on the screen. The image is shortly after inverted in its black and white values (somehow recovering some hope for light and life) and then slowly, gradually fades away.

In the following sections I will describe some interactive installations of mine which use found tv footage and sound.

Figure 2: Memory of the surface, 1992 – 1998

2.1. Broadway - 2000

Broadway is an interactive sound installation. There is nothing to be seen but a free square space of three meters, visually delimited by a strip on the ground. People move freely inside this virtual space. *Broadway* evokes distant places through sounds. Real, imaginary, possible, fantastic, probable, sawn, narrated – a single super-place described by sounds and noises. Moving through the space the visitor gives life to the installation which otherwise remains silent and anonymous. He/she rebuilds an intimate space by pasting perceived sounds and private memories.

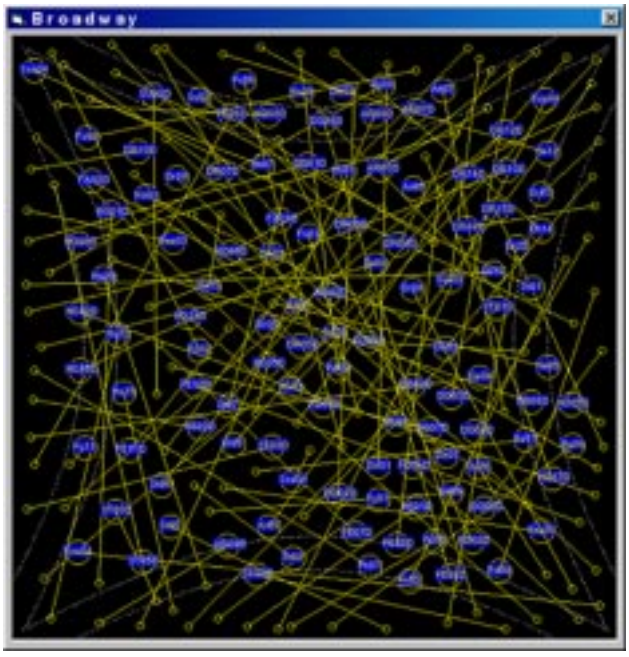


Figure 3: Graphical interface of Broadway's software, 1992 – 1998

Technically the real space corresponds to a virtual space residing in the memory of the computer. The visitor's movement are constantly monitored by a video-camera. As soon as he/she steps into an active area, a previously memorised and spatialised sound is to be heard. The sounds have been gathered on the Internet and in radio and tv programs.

It is not indispensable to see, as far as you can move and recognise possible obstacles. You have to be able to use your internal ability to see, to use your mind's eye in order to combine your personal images to the heard sounds – In contrast to the common ocular sight. Every sound recovers some images' fragment in the visitor's mind (possibly collected watching tv), that gives individual sense to the sound itself: pixel to pixel.

Broadway is a work about memory. It creates an autonomous territory, which comes through each time a visitor walks by, thanks to his/her interaction – that is a physical interaction as well as an interaction that plays on a memory level: combining his/hers and the computer's memory.

2.2. BornBlind - 2000

BornBlind uses the same technical equipment of *Broadway*. It is dedicated to those women who've been raped in Pristina by the Serbs. While the sounds in *Broadway* are set randomly, here they respect a centripetal organisation, like in Mediterranean cities, where the centre is more important than its surroundings. The more one gets closer to the centre of the installation the more one gets closer to the problem. In the virtual space's perimeter there are the sounds captured on the streets in Pristina today, then voices talking about the overall situation (like the general women's condition in that society), while in the virtual space's centre there are the voices of people who are involved in the problem of the rapes and its consequences for mothers and children. No victim is to be heard directly: it's like a black a hole, their voices are missing.

(sounds and voices are captured from interviews made by the journalist Seyed Farian Sabahi)

2.3. Viaggi (Journeys) - 2001

Some years ago, I can't remember when exactly, a group of Mexicans trying to leave their country in order to get inside the Usa, were found dead inside a railroad van, which had been sealed. A similar destiny have the protagonists of a book by a Palestinian writer, Ghassan Kanafani, *Men in the sun*, which tells the story of three men looking for work outside their country: they try to illegally reach the rich Kuwait and find instead a horrible death inside an empty red-hot water cistern under the desert's sun.

These stories have guided me in the *Viaggi (Jouneys)* project, which was meant to deal with "work" as a main theme.

I decided to focus on the difficulties of the search for work, on the moves that have left their mark on generations of people, on entire regions and nations. I've collected and invented the sounds of these journeys, using these as the only means to evoke nightly crossings, illegal embarkations, the violence, anguish and fear mixed with hope of the migrants.

Viaggi is a hostile and claustrophobic territory made exclusively of sounds which the user explores by moving around. It collects screams, moans, sharp orders, reassuring words of a mother to her child (all spoken in different languages), and the sounds of waters

and streets. The installation doesn't want to be didactic. It freezes the moment where the tension is at its height, at the crossing of the border, when two or more different cultures get in touch with one another.

2.4. UnderAttack - 2001

UnderAttack is an interactive video installation dedicated to the WTC attacks. It collects 9 videos documenting the airplanes hitting the towers. The only editing I've done is a slight change of the chromatic values. The sound has been switched off. The work is hosted in a small box which contains a flat screen with the computer hidden behind it and a proximity sensor. Whenever the user moves closer to the monitor, the video plays forward; if the user steps back the video plays backwards too. If the user halts for more than a few seconds, the installation switches to the following video.

UnderAttack is not a work about terror. It is a work dealing with the relationship between the viewer and tv information, about the perception of a mediated reality, about the watcher's obsessive voyeuristic attitude which is protected by the screen – the latter assuring a safety distance between him/her and the real world. It questions the responsibility of the spectator.

Figure 4: UnderAttack, 2001

2.5. LipstickJoystick - 2002

LipstickJoystick is an interactive video installation which gathers a vast number of fragments extracted from cartoons like *Tom & Jerry*. All fragments have been chosen for their extremely violent content, which we do no more perceive as harming. The installation is made of a box with a flat screen and the computer behind it (like *UnderAttack*), but instead of the proximity sensor there is a little joystick through which the user can interact with the cartoons' fragments.

2.6. SuckerToday - 2002

SuckerToday has the shape of a labyrinth, it is a metaphor of a city, of whatever place inhabited by people and by their dialectic thinking and speaking out different cultures and values, in a way which sometimes might be very aggressive.

Figure 5: LipstickJoystick, 2002

SuckerToday is a personal rendering of two violent days in the Italian city Genua during the G8 summit in 2001. While most people where peacefully demonstrating on the streets, there were some very violent little groups, the so called Black Blocks, who went around damaging and destroying. The police reacted by charging the multitudes of those who were peacefully marching. Tension grew heavily and culminated on friday the 20th of July with the death of Carlo Giuliani, a young man killed by the police with a gun.

Hundreds of stills captured from tv programs have been used to document these facts inside the labyrinth.

SuckerToday is meant to underline the affinities between the police and the Black Blocks: both black dressed, with helmets or headgear impeding face recognition, both carrying sticks or truncheons, using rude slogans, moving through the urban space with similar rituals, looking for a physical clash through aggression.

While virtually moving around in the digital landscape of the labyrinth (by means of a joystick, like in a videogame), the visitor activates sounds hidden in the streets: they are fragments from sound works by many artists, poets, musicians which I have collected in the Internet. There is no direct relation with what we see, but a strange and perturbing affinity emerges: maybe because of both the images and the sounds being somehow artificial and excessive.

Technical realization

SuckerToday is a digital labyrinth realised with a proprietary software called Spray, which can manage images, texts, sounds and the interactive movement by the user inside the labyrinth through different input devices: mouse, joystick or an external ultrasound system.

In this case the labyrinth is made of about 400 walls which have been designed with a Cad program, and which are inscribed in a closed space (otherwise you might get lost). On each wall a still captured from television programs has been pasted, while the ground hosts about fifty sound fragments (wav), which are not visible and are activated by the user whenever he/she passes.

SuckerToday, 2002

3. CONCLUSIONS

In my work I recurrently use the following strategies: extracting fragments of sounds *or* images from the never-ending flow of sounds *and* images we (as an audience) are immersed in – these fragments are usually very small and coherent; combining unrelated sound and images; creating a new neutral context for these visual or audio fragments; using simple game-like dynamics and setting (e.g., exploration of a virtual space through the body movements or a joystick).

Through these strategies I want to induce a critical reconsideration both of our role as spectators and of stories or images we usually are familiar with, very often related to dramatic events of our present days. It is my aim never to be assertive or to present a one way interpretation: on the contrary, the user is urged to bring out his memories, feelings, thoughts (in a very private way, though) in order to get an understanding of what is happening in the installation. It is the user through his/her physical and intellectual interaction that gives sense to the installations which otherwise stand still.

The fragmentation of familiar information has a time freezing and estrangement effect that, together with the new neutral context in which the information is set, helps us look at known realities with different eyes. I try to foster the user's active participation through easily understandable, sometimes playful means of interaction.

Another very important aspect is the narrative structure which I try to give to my works: not a tight one but always a legible one. Narrative is something games and media share. And it's it that ideally holds together my installations and that encourages the user's imagination.

A Complete System Of Tridimensional Graphical Representation Of Information : "Crystal Hy-Map™"

David Bihanic

University of Paris I Panthéon-Sorbonne,
CERAP Laboratory (Plastic Arts Study and Research Centre),
27-31 rue Lombard, 92260 Fontenay

root@fxdesignstudio.com

ABSTRACT

This document is a review of a recent project envisioning to implement a system breaking, once and for all, away from the traditional WIMP model and offering a better spatial hierarchical organisation of metadata. For this, it attempts quite particularly to create new description methods in regards to constraints of research space as well as to develop processes of information type assignment, both static and dynamic. Furthermore, it aims to elaborate original interfacing procedures dedicated to semi-immersive interactive display of hypermedia data.

Keywords

Semi-immersive interactive display, Informative space, Semantic analysis, Instinctive interface.

1. INTRODUCTION

Recent research projects carried out on the subject of data processing within a virtual context clearly show that the process that leads the user to adopt a certain strategy rather than another, primarily in the case of numerous situations where complex problems need to be solved, represents one of the fundamental parameters in the development of active principles on which Human-System interfaces depend [4]. These interfaces must allow both to communicate with the different system components and to ensure the control of the decision program in the most uniformal manner. Consequently, a development study is necessary in order to highlight the essential role of man in its operation, a role which necessarily assumes a great level of adaptability of representation and display methods for pertinent information.

Currently, there are numerous visual systems capable of significantly improving the management and analysis of information through well-known gathering and processing procedures [9]. They allow both the understanding of certain common situations and a real anticipation of possible actions and trends.

First published at COSIGN-2003,
09– 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

However, these applications do not offer a great deal of perceptually rich information. They are neither capable of developing a complete spatial representation of information nor of favouring the implementation of new knowledge in complex environments.

Thus, it is of primordial importance to envisage new representation paradigms whose aim is to reach a level of communicational performance of clearly superior systems so as to allow the use of an extended metadata display, an intuitive interfacing system by direct handling as well as a procedure for the localisation and retrieval of multimodal information efficient in big corpora.

That is precisely what this software development project, "Crystal hy-map", aims to do, by attempting to create a real tridimensional graphical representation to explore information. It implements a conceptual approach, resulting from the deformation of the object, which allows to correct the occlusion phenomenon and to build a spatial reality that would be close to the human perceptive system and therefore interpreted favourably as the network of links becomes more complex. In addition, it pioneers new methods for the distribution of information for multidimensional integration spaces (non-hierarchical organisation of metadata, use of original constraints related to positioning, etc.) as well as new procedures for the retrieval of knowledge from hypermedia data [3].

2. WORK HYPOTHESIS

This research project relies on the triple hypothesis that it is possible to:

- convey the structure of the information space through an interactive tridimensional modelling, subject to great constraints such as that of real time
- organise metadata according to their spatial correlation (relation of spatial positioning, closeness)
- imagine an ergonomic browsing space that would allow the user to isolate pertinent information in a precise manner while having an overall view of things.

3. OBJECTIVES

This project questions the current state of the art of displaying information and strives to propose new solutions for the representation of implicit information data.

It turns on five major axes:

- Display of the browsing structure
- Display of non-hierarchical data
- Display of complex operations on the client-system
- Display of multimodal information
- Display of multivaried data.

Development of the software environment relies therefore on certain research works dealing with the combination of display modules, the implementation of an object-oriented approach for tridimensional display, the creation of original datamining processes and procedures for the retrieval of knowledge from data, etc. [6].

4. OPERATING PRINCIPLES

“Crystal hy-map” is a complete representation system that uses a hypermedia structure designed to support a large network of links. It relies on a connectionist system capable of handling numerous data units where information is organised in directories and indexed by geo-semantic analysis. For this, the information is based on a hypermedia browsing repository structured from a “grid” (C.S. : X, Y, Z) for semantic positioning of information revealing new integration procedures mainly between various documentation models and ergonomic knowledge [2]. The main objective lies on the possibility of optimising parameters which constitute the workload for the human operator (e.g. number of objects handled, temporal constraints, parallelism of tasks, length of procedures, etc.) and setting-up a method to diagnose the cognitive workload for the interfaces [11].

Its aim is to highlight the obsolescense of traditional browsing and graphical search tools by insisting on the need to offer new strategies for the processing of information in virtual contexts that are suitable for complex situations. This leads to the implementation of new paradigms for prospective representation of space that make the elaboration of procedures for distribution and project evaluation useful.

Thus, it strives to offer, by way of an interactive feature for semantic distribution of information ("information space mapping"), new methods for the management and handling of nodes, links and annotations ("semantic network") as well as new processes to discover specific knowledge [10].

4.1 Design of an explorer hypermedia

Hypermedia must have organised knowledge on the subject and a modelling of the reading task. It is therefore necessary to ensure a good distribution of skills between the user and the machine and to garantee a true integration of both agents (HMD) in the decision process.

To do so, this project defines a visual organisation concept that would allow to observe and understand information more easily by way of a dynamic tridimensional cartographic display [12].

The original aspect of such an approach lies, beyond formal changes, on the fundamental modification both in terms of distribution, representation, use and production of specialised[*] information and in the elaboration of a new logic to help in the understanding of spatio-temporal processes [7].

The software design process involves the following aspects:

- usability (learning of functionalities only)
- representation of knowledge without logical, hierarchical or ensemblist constraints
- formalism of task description
- multimodal interaction
- vague goals.

Consequently, hypermedia are composed of two reading levels, that of the node and that of the network. In the first level, new browsing functions as well as new data structures are defined. There, information is decomposed into small elementary units that can act as the entry point ("ingress node") or the exit point ("egress node") for the subject. These units, which represent a unique idea and concept, are both syntactically and semantically discrete. They contain several types of information:

- the name that identifies the node
- its interaction data, types of media on which the user may work
- its browsing data, data from which the node function gets its information
- its informative data, data that the user gets via the function.

The second level offers the possibility of browsing the database in a non-predetermined order and of carrying out associative searches.

This idea presents several advantages. First of all, the adding of new data can be aknowledged immediately. Secondly, it is not necessary to plan in advance the way in which the different media will be arranged in the hyperbase, etc.

The latter has therefore no need to be structured or even indexed as what will garantee the pertinence of information is first, the preprocessing of visual and textual descriptors and second, the semantic analysis of information and its location according to areas of influence. Thus, the idea is to consider bases of a large scale and to organise them in such a way that inquiries on flows may be tended to with a transmission rate of satisfying quality.

[*] Work on the notion of space by Klob, Moulthrop, Rosemberg etc.

The description schema will then be adapted to documents and their descriptions.

Raj Jain [8] suggests 3 types of characteristics that are indeed satisfying:

- type *F*u contains "meta-characteristics" that can be automatically retrived from information associated to documents (size, author, format, etc.)
- type *F*d contains characteristics that are directly retrived (derived) from documents at the moment they are entered into the base
- type *F*c contains caracteristics that are also retrieved from documents, but that are only done so at the useful moment.

4.2 "Polyhedral computation"

"Polyhedral computation" is a calculation on the evaluation and distribution of areas of semantic influence of one or several distant "objects" (cf. contextual semantics).

This calculation is the keystone of the system for tridimensional representation of information. It allows for a perfect semantic distribution of information on the object and pioneers, by benefitting from computers' processing rate and graphical capacity, in a new approach for locating and retrieving information for big corpora as well as for searching documents in non-structured bases.

It is organised on an algorythm for geo-semantic location of information. The goal is to "encircle" the object in order to control all its deformations and allow for real time handling of metadata.

5. DEVELOPMENT PERSPECTIVE

So far, projects that have been undertaken have more directly dealt with the elaboration of the system (data display, environment for distributed parallel programming, etc.) as well as database management (modelling of data and multimedia inquiries). The designing of this elaborate operational system leads to the extending of the scope and the concept.

Indeed, "Crystral hy-map" is a very flexible data representation environment that is similar to both the semantic networks and the object orientation paradigm. It is perfectly capable of supporting the implementation of different knowledge representation formalisms [1].

Consequently, one of the short-term goals is to implement this structure in a diagnosis aid interactive system containing a hyperbase, a model base and a knowledge base.

5.1 Prototype for a knowledge base processing system (an expert system for aid in medical diagnosis)

The implementation of dedicated hypermedia, or in other words, the creation of autonomous knowledge entities, does not imply major difficulties. What represents a greater challenge is precisely the ability of making information search possible in a large hypermedia base from an external multi-inquiry processing process.

To accomplish this, we have decided to design an expert aid system for medical diagnosis. The idea is to focus particularly on the diagnosis and problem solving approach while allowing database queries for detailed cases in the hope of improving the doctor's smart workstation [5]. Inference processes (rule detection, selection and application of a learned rule) must be able to handle the data found in the different bases.

Thus, we will probably have to consider a multimedia database – it will rely on a real object formalism or an object layer added onto a relational kernel – for which the user interface would recover the hypermedia concepts developed by "Crystal hy-map".

6. CONCLUSION

Ultimately, through powerful graphic design capable of structuring metadata, "Crystal hy-map" pioneers in a new Human-Machine interfacing concept that facilitates the formulation of inquiries, information display, and surfing within hypermedia.

The development of this system is conditioned by that of research devoted to task description, modelling of cognitive processes involved, by surfing and associating it to other projects such as that of decision aid. Therefore, it integrates the different stages that lead to the knowledge discovery process:

- collection and filtering
- preprocessing
- exploration
- display
- interface.

7. REFERENCES

[1] Agosti M., Crestani F., Melucci M. (1996). Design and implementation of a tool for the automatic construction of hypertexts for information retrieval. *Information Processing & Management, 32(4),* 459-476.

[2] Arents H., Bogaerts W. (1993). Concept-based retrieval of hypermedia information : from term indexing to semantic hyperindexing. *Information Processing & Management, 29(3),* 373-386.

[3] Boyle C., Snell J. (1990). Intelligent navigation for semistructured hypertext documents. *Hypertext : State of the Art, McAleese R. & Green C. (eds.), University of Aberdeen, Intellect, Oxford, England,* 28-42.

[4] Duffy T.M., Knuth R.A. (1990). Hypermedia and instruction : where is the match ?. Jonassen D. and Mandl H. (eds.). Designing Hypermedia for learning, Heidelberg, Springer-Verlag.

[5] Frisse M.E. (1988). Searching for information in an hypertext medical handbook. *Communication of the ACM, 31(7),* 880-886.

[6] Hascoët-Zizi M., Ahlberg C., Korfhage R. & al. (1996). Where is information visualization technology going ?. Pages 75-77. Retrieved June 15, 2002 from LRI (Laboratoire de Recherche en Informatique) database on the World Wide Web : <http://www-ihm.lri.fr/~mountaz/biblio.html>

[7] Hascoët-Zizi M., Beaudouin-Lafon M. (1995). Hypermedia exploration with interactive dynamic maps. *International Journal on Human Computer Interaction, Vol. 43.*

[8] Jain R. (1996). Multimedia System and Techniques. Kluwer Academic Publishers, Norwel, MA, USA, Chapter InfoScopes, 217-254.

[9] Marchionini G. & al. (1991). A study of user interaction with information retrieval interfaces : progress report. *The Canadian Journal of Information Science, 16(4),* 42-59.

[10] Moreira J., Ribeiro C., Saglio J.M. (1999, Avril). Representation and Manipulation of Moving Points : An Extended Data Model for Location Estimation. *Cartography and Geographic Information Systems (CaGIS), N° 2, Vol. 26.*

[11] Thüring M., Hannemann A., Haake J. (1995). Hypermedia and Cognition : Designing for Comprehension. *Communication of the ACM, 38(8),* 57-66.

[12] Wandersee J. (1990). Concept mapping and the cartography of cognition. *Journal of research in science teaching, 27(10),* 923-936.

Signs From A Strange Planet: Role Play And Social Performance In *Anarchy Online*

Andrew Burn & Diane Carr
Institute of Education
University of London
a.burn@ioe.ac.uk
d.carr@ioe.ac.uk

ABSTRACT
In this paper we examine the online multi-player RPG *Anarchy Online*, using social semiotics. Social semiotics emphasizes context, discourse and motivation in sign making and sign interpretation. Drawing on our own experiences of leveling up in *Anarchy Online*, player interviews, and recorded game sessions, we suggest three inter-related and broad categories of motivation: representational, ludic and communal.

Keywords
RPG, role-play, avatars, social semiotics, motivation.

1. Introduction
Anarchy Online (Funcom) is a science fiction styled multiplayer online role-playing game set on the mysterious planet of Rubi-Ka. The population of Rubi-Ka undertake missions, dodge dangerous animals and augment their bodies with nano-implants powered by 'notum', a rare and precious mineral. As you would expect with a role-playing game, each player constructs his or her avatar according to a set of templates relating to species, skills, looks and profession. These avatars are the sign of the player's presence in this fictional world. Because *Anarchy Online* is an online game, the world of Rubi-Ka, with its factions, cityscapes and deserts, is a shared space. Players interact (with varying degrees of skill, civility, hostility or ineptitude) in real time, thanks to their colourful digital representatives, and an in-game chat window.

Anarchy Online (www.anarchyonline.com) is undeniably multimodal, meaning that users re-act to, and act within, the game's world while responding to written, audio and pictorial information. The player clicks on a mouse and keyboard while gazing at an online graphically rendered world. They manipulate their avatar, and type/chat live to other players. In the background a rainstorm or a moody score lend atmosphere. The theory we have adopted in order to make sense of this unpredictable and massive text is social semiotics. Social semiotics (Hodge and Kress 1988, Kress and van Leeuwen 2001; Halliday, 1978) is a branch of semiotics that proposes that the relationship between sign and signifier is socially motivated. Social semiotics places an emphasis on the creative work of the signmmaker, and their transformative use of the available semiotic resources. A further development of social semiotic theory proposes that the landscape of contemporary communication is *multimodal* (Kress and van Leeuwen, 2001). Social actors and communicators reach for the semiotic modes, tools and media most suited to their purpose, and integrate them in ways provisionally determined by generic conventions, but always transformed in use, according to the needs of the signmaker. We are interested in how *Anarchy Online* offers a rich array of semiotic resources to the player, who learns how to deploy these resources for digital dressing up, exploration, self expression and combat. We will also explore how the players' speech-like mode of 'chat' shapes the ways in which roles are played out, and facilitates the players' interaction with each other.

This approach provides us with a frame through which to delineate and conceptualise the repertoire of modes and acts at the disposal of *Anarchy Online* players. While console games might be played with friends, and while much computer game play is contextualised by shared cultural activities of some description (online fan culture, walkthroughs, cheat sharing etc) an online multiplayer game like *Anarchy Online*, is unavoidably public. The graphic world is shared with thousands of other players. Social semiotics is appropriate for our inquiry as it conceptualises signs as arising from, and constructing, social discourse. If the sign making and sign reading activities are discursive and contextual, motivated rather than arbitrary, the initial question becomes: what are these motivations?

While we accept that the answer to this question might well vary from player to player, we propose the following broad (and provisionally titled) areas within which to explore the presence of motivated sign making and sign reading in this game. First, we suggest various **representational motivations** – this category involves presentational, dramatic, narrative and performative aspects within the game. Second, we suggest a **ludic motivation**: an interest in the skills, rules, competition and dynamic engagement invited by the game. This category is named for 'ludology' a term popularised by theorists such as Gonzalo Frasca (www.ludology.org). Finally we explore related categories of **communal** motivations. These involve the game's generic identity, fan cultures, wider digital culture and the taste communities in which it is inter-textually embedded. The notion of the 'communal' is intended to refer to both the social, shared nature of the game, and the sense that the game itself is located within a generic community that encompasses similarly themed fiction and other computer games.

First published at COSIGN-2003,
09– 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

In order to address these questions, our first task was to learn to play *Anarchy Online* ourselves. While we have to date dedicated many hours to this end, we still regularly find ourselves lost and baffled on Rubi-Ka. In addition to self-reporting our own somewhat fumbled forays into this online world, we enlisted three teen-age volunteers to be our informants. We interviewed them prior to their introduction to *Anarchy Online* about their expectations regarding the game, and kept in contact with them as they familiarised themselves with the game world. Over time we conducted further interviews on video, via email, and 'as avatars' within the game world itself. This presentation is an abridgement of a chapter length paper written as part of our project 'Textuality in video games: interactivity, narrative space and role play'. Our chapter on *Anarchy Online* draws on other work we have done over the course of this project, including Andrew's work on multimodality and social semiotics in the *Harry Potter* and *Final Fantasy* games (2002, 2003), and Diane's work on genre, and narrative and ludic discourses in RPGs (2003a,b). For the sake of brevity, in this presentation we will limit our discussion to the various motivations we have discerned, rather than addressing the wider implications of performance or identity.

Figure One: Aisea applauding

***Anarchy Online* (copyright Funcom)**

1.1 Representational motivations : "Welcome to Rubi-Ka"

Anarchy Online is a role-playing game (RPG), which means that it owes certain generic factors and a rules system to table-top role playing games such as *Dungeons and Dragons* (TSR 1980s). Whereas console games generally offer the player a ready made avatar with which to steer their way through the game (such as Lara Croft of *Tomb Raider* fame), RPGs typically offer the player the opportunity to construct their own avatar/protagonist utilising a set of templates relating to profession, species, and physicality. In one sense players of *Anarchy Online* are choosing options from a limited set of paradigms in order to construct the 'syntagmatic bundle', which, once named, becomes their avatar. But, in actuality, even the few options that relate specifically to the avatar's most overt physical characteristics (species and gender) are subject to a multiplicative effect that produces considerable variety. The player selects from four very different species, each with different strengths. The player then selects a gender. Each of these choices will have ramifications for the subsequent options that the player is presented with. The player selects a face for their avatar, and then a height and a body weight. Variation is further multiplied when the player selects from one of twelve professions (which determines a character's initial wardrobe, among other things). The player is then invited to give their character a nickname, which has to be unique. Then the player decides on the political alignment (neutral, clan or corporate) of their avatar. These options relate to the game's back-story and have geopolitical consequences for an avatar.

This simple set of options produces thousands of possible characters. Though this is still a restricted set of semiotic possibilities, it does result in real diversity within the game, and a sense that each avatar is unique. Of course variety and distinctiveness are further developed during play. Players can buy clothes (from party dresses, to protective hoods and boots) and win armour during missions, thus manipulating and individualising their avatar's appearance further. Players also choose the skills and weapons acquired by their avatars as they 'level up'.

1.2 Character generation: Nirvano, Grayse and Aisea

Anarchy Online dramatises the character selection process. A new game opens with a view from orbit, and the player is welcomed to the "territorial space" of Rubi-Ka. The 'camera' then swoops along the corridors of a space station, and the player is invited to begin DNA sequencing a body for their new life on the planet below.

Andrew: During the construction of my first character, I felt compelled to make him bald and a little overweight, as if there was a kind of honesty in resisting any temptation to construct a muscle-bound representation. Other choices, too, felt unexpectedly loaded: it is not possible to choose gender, for instance, without, in some basic way, saying that this character is going to share a set of potential cultural dispositions with me; or I'm going to adopt a set of dispositions profoundly unfamiliar (our three player-interviewees all expected to choose avatars like themselves in some respects). Nirvano is a solitus, the species on Rubi-Ka that is closest to human. My character is called Nirvano, a masculinized version of the name of an American student I once taught briefly; it's reminiscent of appealing Buddhist characteristics; and at the same time appropriate to the sci-fi setting of the game. The choice of a name is in itself an intensive semiotic activity – loaded with associative signification. It may be witty, misleading, serious, erudite, genre-based. Nirvano is a martial artist. I had begun with a liberal disinclination to carry weapons, and initially tried to play as a 'meta physician' character, but I discovered that having a character survive, even in the training level, required that they be able to fight off aggressive small animals.

In some respects, this character generation process is like the offstage space Goffman (1959) terms the 'back region', as distinct from the front region where the performance of role is taken out into the world. The back region is the more private

space, where the role is prepared. In life, for instance, it may the private domestic domain, where the elements of the public role are assembled for the day. In the theatre, it is the dressing room. In the game, it is this offstage laboratory like space where the parts of the avatar are laid out. However, there is another space that also corresponds to Goffman's back region – the world on the other side of the screen, where the player sits at a computer, entering the gameworld through the interface technologies of keyboard, mouse and screen. In this space, while creating Nirvano and stumbling with him into the training ground with other new players ('newbies') I muttered to myself, exclaimed aloud when surprised (or killed), jotted down possible names when the nicknaming process was in train (the first four names were rejected). As I think of this, I am reminded of the *Numbskulls*, a comic strip in which the central characters were little creatures occupying the brain of a man, who they operated by a series of mechanical devices. The construction of character felt, at least to me, like a set of parts for a character, a resource for dressing-up, a character kit. The character might be 'starting from scratch' but just like when my 6-year-old nephew dresses up as Batman, you know that you are supposed to behave in a particular kind of way.

Diane: While I have played in the guise of a short, bald engineer with a moustache, called Festa, the avatars that I have spent most time playing are female. The first character I built was (and is) called Grayse. I choose to play as a doctor, believing that her ability to heal other players would help facilitate friendly interaction and team play. Grayse is small, dark, and named after a friend. Grayse is a nanomage- a small indigenous humanoid with certain genetic and technological advantages. Generally in RPGs magic users become powerful over time, but initially they are vulnerable. This, it quickly became apparent, was also true of technology dependants on Rubi-Ka (*Anarchy Online* is a science fiction and the alien technologies operate much as magic would in a fantasy RPG). After being killed numerous times by toxic rodents, I put Grayse to one side in order to learn the basics of the game with an easier and stronger character, and returned to the character generation phase. Second time around, I built a martial artist named Aisea. Aisea is also named for an old friend. Players can run up to seven characters in *Anarchy Online* (although they can only operate one at a time within the game world). Aisea is, like Grayse, small dark and female. She is an Opifex, so she is genetically predisposed to absorb the power of 'notum'. I choose to play a martial artist because I like martial arts movies, and because it suggests to me more interesting combat possibilities than shooting does. An additional incentive was that I had noticed other female martial artists in the game wearing cocktail dresses. The idea of combining lethal moves with pretty clothes is appealing. As Aisea has levelled up she has learnt some relatively spectacular moves, including flying kicks. Aisea is tiny compared to many of the other characters, and her smallness gives me pleasure: it makes her look deceptively harmless, and it signifies a very satisfying combination of precision and violence.

Aisea and Nirvano are partial representations of us, at least to the extent that they express various preferences. Each character is also bearing generic markers (they are martial artists in a science fictive space). We have chosen from a restricted menu of semiotic resources. Our selections are down to personal prerogative, but the menu of choices on offer is afforded by the game. Thanks to the multiplicative nature of the options on offer during character creation, our avatars are visibly distinctive. Our avatar's names are unique as well: nicknames are entered using text, and the only restriction is that is singular.

Figure Two: Aisea fighting Highvoltage the TechnoSlave
***Anarchy Online* (copyright Funcom)**

1.3 Newbie to veteran; templates to biographies

When we joined the other players in the training grounds and cities of Rubi-Ka, it became apparent that the seemingly incidental choices we had made in the privacy of the character generation chamber would have repercussions. Andrew, for example, felt bound to construct an 'honest' rather than a wishful persona, and as a consequence, found himself in the surreal position of taking another player's derogatory comments about Nirvano's fitness personally. Over time it became clear that the private frameworks through which we assembled our avatars, using the resources supplied by the game, were more revealing and autobiographic than we had initially appreciated. Our choices related to how we feel (however vaguely or unconsciously) about managing shared spaces in real life. It is very probable that this is indicative of our lack of experience with online invented personas.

Perusing the player's forum makes it clear that many experienced players make very specific 'role play' choices, right from the start. They have intentions to play a type of character, with a particular background and allegiances. While we made our early selections based, seemingly, on judgements that refer back to our sense of our real selves, many veteran players make their selections based on a specific fictional identity that they have (partially at least) already designed. Our early attempts at character generation were very much, as Andrew has described them, a kind of playful dressing up. What became clear only later, was that our playful choices were more indicative or revealing, and less spontaneous and whimsical than we had

thought. Returning to the character generation stage to invent a new character after you have played the game (even if still a novice) is a different experience. Once you have played, your familiarity with the game contextualises the offered resources and the repertoire of potentials offered by the game connote their eventual expression in the game world.

A percentage of players of *Anarchy Online* are committed to Role Playing. Role Players (or RPers) invent characters with biographies and histories that far exceed the templates on offer. The intention is that these characters partake in shared events and improvised scenarios. There is an active Role Player forum on the *Anarchy Online* website, where players meet to discuss the state of the game. A few frequent topics of discussion include:

- The relationship between Role Players and the general player population
- The best way to begin role playing, as opposed to just playing the game
- How to spot a Role Player
- The right way to play a 'baddie'.

Vixentrox (dec 31, 2002) suggests that when creating a character "a brief background outline is a good place to start" and "If you have multiple characters…make sure they RP as different people. My main character has a 'step sister'. They trade insults and don't like each other very much. The one is more fun loving…the other is more serious and stern." Its obvious that Vixentrox is referring to traits that have little to do with the templates offered by the game, other than that it's probable that a face could be selected that would, at least for that player, communicate a certain kind of personality. Lillemjau (Jan 5th 2003) replies to a beginner's request for Role Playing advice with the following:

"I want to address your character development…in between the background history and personality traits, add some good and bad habits, strengths and weaknesses things your character loves and hates. Those little things makes the depth of him/her more interesting".

As Lillemjau's and Vixentrox's posts to the player forums indicate, for commited role players the character templates offered by the game as a set of resources are only a set of starting positions. The manner in which a player might interpret and then perform the identity of their avatar exceeds their indexed attributes. In this sense, the resources offered by the game are what Kress and van Leeuwen (2001), following Halliday, (1985) call systems of *meaning potential*. In a game, resources and rule-systems are offered – but like language, there are not necessarily limits on what can be "said" using these resources: in this case, the construction of complex, highly developed dramatic roles.

The player makes choices about the 'look' of an avatar, and the style of play that they undertake. We tend to think of this category of sign making as being motivated by expressive, narrative and dramatic concerns. What typifies these motivations is that they are not limited by the explicit ludic imperatives of the game (goals and scores, for example). For motivations that more directly concern the game structure, and the real time events of play, we turn to what we have called 'ludic motivations'.

2. Ludic motivations : game, goals, strategies

As has been pointed out by various videogame theorists (Juul 2001, Eskelinen 2001) narrative discourse contains prior events that are ordered in time (or plotted) and related to the user, whereas playing a game involves events that are ordered, at least in part, by the player, and that unfold in the real time of the user. This differentiation is central to the distinction between our first two categories of motivation. The ludic qualities of *Anarchy Online* are those parts of the game that make it a game: strategy, goals, real time events, chance, rules, skills acquisition, exploration and levelling up. The narrative and representational concerns discussed above rely on schemata that are not necessarily made explicit on screen. A character's biography, for example, will refer to previous events (as opposed to the real time events of play). One immediate clarification needs to be made: the borders between these ludic and the narrative strata are not particularly distinct. There is some discussion on the player's forums as to what, for example, distinguishes role play (play with a self professed narrative agenda, as described above) from general play, where players go on missions or explore the game world using their avatar as a game-tool, rather than as a character per se.

Figure 3: *Anarchy Online*, copyright: Funcom. Ludic motivations: Image of the character 'Japhis' with her statistics screen

Ludic modes foreground the role of the avatar as game component, as symbolic unit of strategic value (like a chess-piece). Accordingly this category revolves around considerations of 'how to play'. Ludic activities on *Anarchy Online* include going on missions, selecting one style of 'profession' over another based on your preferred style of strategising (sniper over martial artist, for example) and directing energies towards the accumulation of experience points (through goal attainment) that enable avatars to 'level up'. In terms of sign making and sign reading, much of this activity is focused on the operations of the player's avatar.

While expert players traverse the game world, beginners struggle to move their avatars at all. It's difficult to see what you need to see, and tricky communicating properly with other characters.

Our new avatars staggered around, made false moves, rotated wildly, or were strangely still. When we wanted to quit out of the game, we realised we didn't know how. We selected 'quit' from the menu, and were informed by text that we must sit down before we could quit. But we could not find out how to sit down. Fortunately the training ground is full of other players of varying degrees of expertise. Once you have mastered the ability to type/enter basic conversation (not as simple as it might sound), it is possible to ask advice of your fellow players. There is also an open chat channel just for 'newbies' which tends to be full of questions and requests for aid, including some quite odd ones: "my head is stuck in a wall, can somebody help me?"

Just as the character templates supplied by the game initially appear limited, the various motions of the new avatar appear stiff and mechanistic. They are stilted until the player has gained a certain level of familiarity with the controls. At first the player has to make a considerable conscious effort to drive the avatar through simple actions, such as turning a corner without bumping into a wall, or running along a path without falling over the edge. The commands that are constantly used soon move to a stage of less conscious manipulation. This is an acquired skill, a literacy, a fluency. The avatar moves in the world through the combination of a set of technological potentials for sign-making and the player's skill in deploying those signs, much as you might move through the world in a car, making both a functional journey and a social performance, by exploiting those potentials through learned manipulative skills.

The actions of the avatar depend on the player – and these actions involve the manipulation of technologically mediated signifiers. These signifiers recall what Halliday (1989) has called the *restricted language* of games. His example is contract bridge, where very limited sets of signifiers (such as the four suits in a deck of cards) can multiply with other sets (such as the numbers of such suits which can be bid) within the rule structure of the game. In spite of the restrictions, the range of possible combinations, the ways in which they relate to the rules of the game, and the way all this in turn is determined by strategic collaboration between players, all make for a complex and creative activity requiring considerable skill. In computational linguistics, too, restricted languages have found a new significance as bounded systems which computers can handle easily, as opposed to the unpredictable, unbounded nature of natural language.

The player responds to the game's rules, which are expressed in semiotic terms as missions, weapons, rewards and first-aid kits. These are central to the game's challenges and game-play. The player has available a restricted language of avatar movement – run and walk, directionality, jump, but each movement, however simple, is immediately a more complex act semiotically. This is because the movements are not made against an empty white background: the gameworld contains both a landscape and other characters and creatures, any move of our avatar makes a syntagm – a move forward combines our avatar's move with the landscape; and with other avatars. As the player acquires fluency, the ludic and representational qualities blur: a high level avatar will be visually distinctive, personalised by exotic armour and monstrous pets.

If we decide to take our avatars for a walk outside the city gates of Borealis, we 'walk forward' and immediately other signs in the 3 D multimodal world collect around this simple action – we can "see" a new landscape outside the city; the music changes; the sound of our feet on the ground changes to the crunch of a sandy path. Simply wandering around in the gameworld involves the combination of a highly restricted language of avatar action with the much bigger language of the gameworld and its contents. The latter is, in principle, an unrestricted language, as anything can be designed into it, whereas an avatar's movements depend on the game's mechanics and the player's fingers. In semiotic terms, however, it is clear that the 'design' of our avatar's wandering, is a joint activity, with three principal co-designers at work. The player designs a walk through the woods; and the game's programming designs elements around us. The other co-designers of our experience are fellow players, who may affect our experience to a greater or lesser degree. In addition to the representational and ludic motivations we have discussed, these communal motivations (informed by various social and cultural expectations) shape the player's experience of the shared game world.

3. Communal motivations: Sharing Rubi-Ka

We consider communal motivations to include player expectations, genre, trans-textual content, the wider gaming community, role-playing in other games, and fan culture; and the fact that the game is a shared, largely public space.

When we interviewed our three (male, teenage) player-consultants about their expectations of *Anarchy Online*, before they began to play, certain communal motivations were clear – all three of them had played online games before, and all had ideas about the pleasures and options that the game would provide. Their expectations in terms of ludic motivations were clear: they had well-articulated expectations of certain aspects of play, such as the kinds of combat they would be able to engage in, the ways in which they might level up, and the kinds of choices they might make about their avatars which would give them strategic advantages in the game.

Once one of the interviewees, Tim, began to play *Anarchy Online*, he employed a trans-textual approach to selecting his character's name. Interestingly his action associates this game with its *Dungeons and Dragons*/Fantasy roots, rather than with its science fiction setting. Tim found an online English to Elvish translator, and used his own nickname to generate a name for his avatar, as he explains in this e-mail message:

> Name: Belithralith - soulish (my nickname)
> translated into elvish on an internet translator.
> Breed: nanomage - just look kind of misterious
> Gender: male - 'cos that what i am i 'spose
> Profession: Agent - all i can say is: sniper rifles :)

Clearly, then, the semiotic motivation here depends on related discourses – the name is borrowed from Tolkienesque narratives and live action role-play games. Tim explained his predilection for sniper rifles in terms of his past experiences in FPS games. Like our own choices of body, Tim's decisions are also influenced by a sense of connection between his online persona and himself – "'cos that what i am i 'spose". All three interviewees expressed distaste for the 'dressing up' potentials of the game.

It was, perhaps, in terms of the social motivations of the game that our interviewee's expectations were most interesting. They expected quite specific kinds of relations between avatars and players. In particular, in relation to gender, they constructed a modality that selectively 'read through' the appearance of other avatars. The boys were very certain that behind all female avatars, are 'fat American (male) teenagers', which they then quickly revised to 'fat middle-aged American men'. However, when asked what kinds of avatar they would choose, they all said they would be male, human, and as like themselves as possible. This kind of expectation seems rooted in a broad stereotyping (even when females are visible in an online game world, we're invisible!) which we expect to break down to some extent when they actually encounter other players. Their distrust also appeared to be rooted in a discourse of 'internet suspicion', born of an alarmist and wary attitude towards predatory online duplicity. In the case of these three teenagers, such suspicions are reinvented as a 'knowing 'discourse.

One option when playing *Anarchy Online* is to play solo – to refuse to engage with other players. However it is not possible to ignore their existence in the world: they will run past you, hold you up in queues at mission terminals, stand next to you in shops, have conversations you can 'hear', and approach you with questions or requests to join them. As we explored these possibilities, it become apparent that the interaction with other players is channelled primarily through two (occasionally disarticulated) channels, or modes:

- The visual, animated aspect of the avatar (how they act, how they look)
- In-game live chat (typed and entered by players in real time, some 'in character' some 'out of character')

The first visual mode of sign making (and sign reading) is generated by the visual design, actions and animating of the individualised avatars. Our avatars enable us to occupy the gameworld, and to approach others in a way that signals some expectation of reaction. This particular system of *meaning potential* involves the way in which our avatars are equipped with a repertoire of blended animated movements expressive of emotion or certain kinds of social communication. These include waving, nodding, laughing, pointing, dancing, and various other rude, humorous and expressive gestures. When Aisea first appeared next to Nirvano at the appointed time and place, she was waved her arms in vigorous greeting. Andrew immediately became conscious, for the first time, that Nirvano's arms were pinned helplessly to his side; and he had to ask (i.e. type into the chat-box) "How do you do that?" This visual aspect of the avatar employs various potentials: costume, body, face and movement, and all are elaborately expressive. There can be no lapse from this locale, because the player's presence depends on and manifests as the avatar.

In-game live chat involves the typed entries of players, and this mode is comparatively flexible: chat is at times the 'voice' of the avatar, but at other times it's clearly the player who is talking. This speech (which has no actual vocal or audio component, it is typed and read) is entirely at the player's disposal, so that it is possible to construct every shade of commitment to the avatar's identity: to slip in and out of role, to maintain the role at a low level, to modify the role, to speak in your own voice (as a player) from behind the mask, or to speak in the voice of the mask. Game chat swings from 'in character' dialogue, to ludic orientated dialogue about gameplay and team formation, to observations on other avatars' physicality or equipment, to completely unrelated sociable chat ("Hi, where are you from?" or "Are you a real girl?"). This chat mode, unlike the specific, restricted language of the animated emote repertoire, is an unrestricted natural language and, like our nicknames, it is a form of semiotic work that is completely player-produced. The huge majority of in-game conversations that we have witnessed are as direct and abridged as the one we include here. In this sequence, Nirvano is trying to join a clan that is in the process of being formed by Stormthunder, Regrat, Articspider, Thie, Demonbuster and Fithelement:

Stormthunder:	make name
Stormthunder:	first
Nirvano:	yeah
Thie:	fighters of the lost realm
Stormthunder:	I got to go now
Regrat:	athen whompa? [this refers to a portal to another city]
Stormthunder:	hurry up
Articspider:	make me leader
Thie:	ok
Stormthunder:	u can make name
Thie:	no me
Stormthunder:	fith hurry up
Regrat:	thx anyway ☺

Stormthunder, although she expresses no ambition to be leader, or to decide on the name of the group, is assertive in assigning tasks to others, using imperative forms ('make name'; 'hurry up'). Similarly, the other dominant theme of the conversation, the choosing of a name for the guild, is appropriate to role-playing games (even if it appears to have strayed in from an alternate genre, i.e. fantasy). The group eventually agrees on "Fighters of the Lost Realm" (Nirvano's only contribution is to correct the spelling of 'realm' as he can't bear the thought of wandering round belonging to a misspelled guild). The players are clearly employing a ludic rather than narrative mode, organising a team with strategic rather than dramatic motives. The avatars were not talking in a manner dramatically consonant with the visual style of their character. Rather, the players are communicating in the universal, compressed dialogue, of synchronous chat, with typical orthographic and stylistic features; a mode that implies certain cultural and perhaps age-related attributes (a familiarity with online environments, an ability to talk and read 'txt'). Werry (1996) notes several linguistic features of Internet Relay Chat, such as abbreviation, paralinguistic cues, and actions and gestures, each an adaptive strategy to allow it to behave as much like speech as possible. The 'talking' in *Anarchy Online* displays many of the same features. The game produces the addressivity necessary by showing the names of the speakers. Abbreviations are used, both grammatical ('make guild') and orthographic ('u can make name'), while facial expression is simulated with Regrat's smiley. While one of the motivations may be the desire to

replicate the 'feel' of speech within a typed and read mode, its also completely possible that the urge to save time (or "hurry up") is motivating the players.

In this particular example of synchronous chat, the players have a high ludic motivation, and a low commitment to developed roleplay or characterisation. In the above exchange, no one was interested in the character, history, or personality of his or her fictitious characters - only in getting the job done. The dynamic properties of the exchange were not dramatised elements of invented roles, but the real impatience, assertiveness and indecision of the human players. In terms of multimodal theory, the semiotic effect of this kind of role-play is a pulling-apart of the two modes through which the avatar acts – the animated image and the written chat. In this instance, the two modes are only loosely connected, because the dressing-up part of the role and the strategic decision-making are only loosely-connected.

Though we have no space here to develop an account of how the modes of image act and speech act integrate differently where players are committed to Role Play, we can give one brief illustration. Such players will infuse their exchanges with invented personas. In such cases abbreviation is clearly not an issue, as is clear in this exchange witnessed by Andrew in passing: "I bow to your superior wit and wisdom, and withdraw from combat".

4. In conclusion

The broad categories of motivated sign making and sign reading that we have been exploring in the context of *Anarchy Online* do not occur in isolation, on the contrary, they are combined during play. We conclude by relating an in-game encounter intended to illustrate the co-existence of the various modes and motivations. To recap, we described these styles of motivation as:

- **Representational** (dramatic, performative, figurative, graphic, narrative)
- **Ludic** (game orientated: scoring, levelling up, the avatar as tool)
- **Communal** (generic and other expectations, wider online culture, the shared nature of the game world).

In the middle of a play session Nirvano (Andrew) and Grayse (Diane) were trying to decide whether to head straight out on a mission, or go shopping, when we were interrupted. A 'voice' intruded on our conversation (in the form of a line of text) to ask if Nirvano "was pregnant". Andrew's understanding of this was that the new arrival was making an offensive remark about his avatar's girth. Diane's understanding was that the player was mocking Nirvano for hanging around with girls (and so retorted, playground style "why, are you?"). Thus we met Rafayel, an avatar with a 'male model' physique, wearing high heels, thong style underpants and sunglasses.[1]

For Andrew, Rafayel's comment was confronting because Nirvano wears aspects of Andrew's real body image. Diane's (Grayse's) response was also triggered primarily by her 'real world' identity: Rafayel was hassling Nirvano for associating with a female. Rafayel made more jokes about Nirvano's appearance ('i can hear the baby kick") and then compounded the provocation with mildly confronting actions (walking close enough to Nirvano to make contact, and then apparently bouncing off). In both his dialogue, and in his stance, Rafayel completely ignored Grayse. Grayse resorted to conciliatory compliments about Rafayel's shoes, but to no avail. Inspired by the tattoo that Rafayel sports across his chest, Grayse asked about where to get one for herself. Rafayel ignored Grayse. Nirvano repeated the question, and Rafayel answered "on missions mostly", referring to game play (the tattoos are a mission reward).

We responded to the representation in front of us (by noticing the tattoo and his bizarre dress sense). Rafayel initially offended both of us, for completely different reasons. We both assumed, throughout the encounter, that Rafayel was a male player. There is of course, no reason to believe that is the case – 'Rafayel' might well be an adolescent girl or a grandmother, but while in this particular embodiment, he was male to us. In his chat (even when 'out of character') Rafayel presents himself as male. Nirvano and Grayse (and Andrew and Diane) experience him as a male presence, because of the way he looked, the way he acted, and the things that he typed. Nirvano and Rafayel met by accident later that same afternoon and had a friendlier conversation. Rafayel offered Nirvano an in-game object for his in-game apartment (a lava lamp, actually) and typed that he was "from Sweden".

As this encounter demonstrates, while it is possible, in part, to distinguish between the various motivations shaping the reading and making of signs in a massive, shared, graphically rendered world like that of *Anarchy Online*, in practice, these motivations are simultaneous. Just as a deceptively simple set of templates combine to create a huge range of possible avatars, the motivations (representational, ludic, communal) that we have examined all mesh during play, proliferating, compounding and informing one another. In practice these motivations become ambiguous and multiple. The invented persona of the avatar remains curiously shot through with aspects of the everyday. The game's elaborate science fiction locale plays against the abbreviated pragmatics of chat, levelling and team formation, and the available semiotic modes of animation and speech-like chat allow for these ambiguities.

5. References

1. We have changed this avatar's name.

6. Bibliography

[1] Boal, A (1979) *Theatre of the Oppressed*, London: Pluto Press

[2] Burn A and Parker D (in press 2003) '"You Believe in Harry Potter"; The Modality of Computer Games' in *Analysing Multimodal Texts* London Continuum

[3] Burn, A and Schott, G (under review) 'Heavy Hero or Digital Dummy: multimodal player-avatar relations in Final Fantasy 7'

[4] Carr Diane (2003a) 'Play Dead; genre and affect in *Planescape Torment* and *Silent Hill*' in *Game Studies* May 03 vol 3 issue 1 online at: www.gamestudies.org

[5] Carr Diane (2003b publ. pending) 'Character Building; Pleasure Transformation and Flow in *Baldur's Gate*', forthcoming anthology *Gameplay: Pleasures, Engagement, Aesthetics* (Swalwell M and J Wilson, eds).

[6] Eskelinen M (2001) 'The Gaming Situation' *Game Studies* vol 1 issue 1 July 2001 www.gamestudies.org

[7] Goffman, E (1959) *The Presentation of Self in Everyday Life*, New York: Anchor Books

[8] Halliday, MAK (1978) *Language as Social Semiotic: The Social Interpretation of Language and Meaning.* London: Arnold

[9] Halliday, MAK (1985) *An Introduction to Functional Grammar*, London: Arnold

[10] Halliday, MAK (1989) *Spoken and written language*. Oxford: Oxford University Press

[11] Juul J (2001) 'Games telling stories? A brief note on games and narrative' *Game Studies* vol 1 issue 1 July 2001 www.gamestudies.org

[12] Hodge, R and Kress, K (1988) *Social Semiotics* , Cambridge: Polity

[13] Kress, K & Van Leeuwen, T (1996) *Reading Images: the Grammar of Visual Design*, London: Routledge

[14] Kress, G & Van Leeuwen, T (2000) *Multimodal Discourse: The Modes and Media of Contemporary Communication*, London: Arnold

[15] Lemke, J (2002) 'Travels in Hypermodality', *Visual Communication 1*(3): 299-325. 2002.

[16] Mortensen, Torill (2002) 'Playing with Players; Potential methodologies for MUDS' *Games Studies* July 2002 volume 2 issue 1 www.gamestudies.org

[17] Van Leeuwen, T (1999) *Speech, Music, Sound*, London: Macmillan

[18] Werry, C (1996) Linguistic and Interactional Features of Internet Relay Chat, in Herring, S (ed) *Computer-Mediated communication: Linguistic, Social and Cross-Cultural Perspectives*, Amsterdam: John Benjamin

[19] Wright T, E. Boria and P Braidenbach (2002) 'Creative Player Actions in FPS Online Video Games – Playing Counter Strike' *Games Studies* December 2002 vol 2 issue 2 www.gamestudies.org

7. Acknowledgements

This paper was writing as part of the 'Textuality in Video Games' Project at the Institute of Education, University of London in the Centre for the Study of Children Youth and Media, September 2001-August 2003. For further details please see
http://www.ccsonline.org.uk/mediacentre/Research_Projects/main.html

Lexicon

Andy Deck
Artist
Artcontext.org
artcontext@teacher.com

ABSTRACT

Lexicon is an extensible online artwork that invites public collaboration. The work addresses the relationships between verbal and software languages, and between free expression and coercion.

Keywords

Word play. Visual poetics. Programming. Interactivity. Creativity.

1. ENCOUNTERING SOFTWARE

Understood as an interactive experience, Lexicon is a form of software that has to be discovered rather than immediately understood. It offers a series of encounters: with programmatic images, and with strangers who are visiting the piece at the same time. People passing through the Lexicon, are its performers. The patient and the curious may discover that with words they can control the sequence of events, as Lexicon's authors.

1.1 Word and Image

Understood as a tool, Lexicon combines word play with image processing. Each of Lexicon's visual effects and transitions is associated with a word. Arranged in various combinations these word produce visual phenomena that could be described as interactive montage or even, perhaps, as creativity (yours, not mine).

2. CONSTRAINTS

2.1 Beyond Interactive Spaces

> Just as the astronaut broke free of the reality of his native world in landing on the moon, the cybernaut momentarily leaves the reality of mundane space-time and inserts himself into the *cybernetic straitjacket* of the virtual-reality environment control program [1]. – Paul Virilio

Most graphical interactivity has come to resemble the exploration of preconceived spaces. It differs little from video games etched into mazes of ROM (read only memory). Fascinating, perhaps. Targeted at the consumer, these spaces are not conducive to metamorphosis and reorganization. Lexicon deemphasizes spatial navigation, emphasizing language and performance instead. This is not to demonstrate that by eliminating the spatial metaphor in interactivity we will suddenly elude the *cybernetic straitjacket*, but rather, to delve into the problem of public creativity in cyberspace.

2.2 Participation

Lexicon invites participation at a number of levels, including the writing of scripts that affect what others experience when they visit the site. This writing is easier than it may seem. There are special words that cause visual effects to occur when a script is "running"; however, people can make meaningful scripts even before they understand how Lexicon works. This leverages what people already know, and tries to make the learning process less rigid.

2.3 Kinds of Expertise

Lexicon balances the image between the time-honored practices of written narrative and the often frustrating dominance of programming codes in digital media. Collaboratively, participants can intervene either as authors who understand words and their meanings, or as programmers, or via a middle-path that involves a little of both roles. Of course, people can simply click their way through the piece, as spastic apes always do when browsing the Internet. (Actually, call me cynical, this is the path of least resistance and the one I think most people will follow, too. So be it.)

2.4 Authoring Software

At the divide between verbal composition and computer code, Lexicon reveals parallels, possibilities, and significant differences. When communicating with words, humans generate an enormous variety of combinations and meanings. Almost everyone can do it, too. The situation for computer-mediated communication is somewhat different,– especially if images are involved. Visual creativity in interactive media is mired in complicated "authoring" software and programming. Lexicon points graphical interactivity toward the dialogic model of spoken languages, and the uncertainty of shared experience. Like stage plays, the various performances of a Lexicon's scripts can produce diverse results. Lexicon offers a live telematic medium for communication and verbal-visual composition.

3. ILLUSIONS OF CREATIVITY?

However, significant aesthetic biases are more deeply embedded in the software, in the codes that actuate the scripts and translate words into imagery. There are limits to the variations of imagery that can be achieved through changes of sequence and performance. In this regard, Lexicon frames some important

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

questions about the nature of creativity in the context of software. Who will expand the vocabulary?

3.1 Literacy as Commodity

Whereas in speech anyone can coin a neologism without special training, programming visual poetics requires specialized knowledge. Lexicon is designed to allow its vocabulary to grow, yet it does not offer a northwest passage to avoid the expertise problem. Whereas the dominant paradigm of software development treats code as a commodity provided to consumers by experts, Lexicon articulates the linguistic character of software and asks whether something of the old, free-speech paradigm can be salvaged.

3.2 Conclusion

People who are familiar with the Java programming language (or who are simply ambitious) can add to Lexicon's vocabulary using the Lexicon Development Kit (LDK). It is unclear whether this will inspire and enable much participation. Even if it does not, Lexicon will have demonstrated problematic changes in language, authorship and creativity that do not appear to be well understood by the public,– or by artists for that matter. Moreover, it may serve to illustrate concretely the linguistic importance of the open source movement, which, after all, is more than just a vague artsy meme. The Lexicon codes are in fact available for review, revision, and cooperative invention.

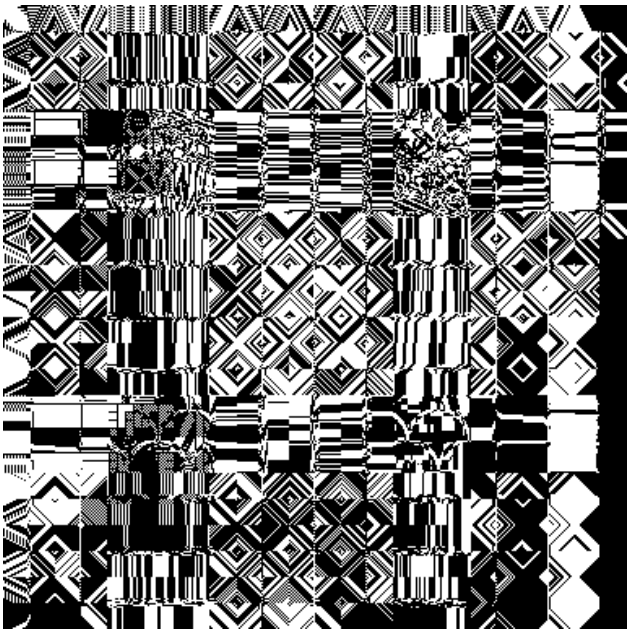


Figure 1. An image from Lexicon, artcontext.org/lexicon/

4. REFERENCES

[1] Virilio, Paul. *La vitesse de libération* . Editions Galilee, 1995.

Synaesthetic Performance In The Club Scene

Annet Dekker
Netherlands Media Art Institute,
Montevideo/Time Based Arts Amsterdam
+31 (0)20 6237101
annet@montevideo.nl

Introduction

"Let's pretend there is a way of getting through into it somehow, Kitty. Let's pretend the glass has gone all soft like gauze, so that we can get through. Why, it's turning into a sort of mist now, I declare! It'll be easy enough to get through--" She was up on the chimneypiece while she said this, though she hardly knew how she had got there. And certainly the glass *was* beginning to melt away, just like a bright silvery mist."

from: *Alice Trough the Looking Glass*, by Lewis Carroll (1871)

In our daily life our various senses are under constant assault. Visitors in shops, cafes and even on the street get overwhelmed with beats accompanied by flickering lights, video and digital imagery, all of them trying to keep up with the music. These developments are most prominent in the club scene where the sounds merge with light, images, smoke and even smell. After the popularity of the Disc-Jockey (DJ) the Video-Jockey (VJ) entered the club scene in the late '80s. The term VJ was introduced in the mid-'80s by television broadcaster MTV. At first it was used for the presenters of their programs, but soon the term was taken over by the people in the club scene for the person in charge of the flow of images that accompanied the music. At first sight the work of a VJ looks similar to that of a video artist. Using the same equipment, it has a similar aesthetics. But in contrast to most video artists the VJ only operates live. In this way VJ shows are related to performance art, especially the live creation of an experience where different senses are being triggered. Most VJ's see themselves having a direct link with the rise of House music in the mid-'80s and the popularity of video clips. In this paper I will show that the now common combination of DJ's and VJ's in clubs is a recent form of synaesthetic performance.

Synaesthesia is the recreation of a sensation through sounds, scents, the visual, the physical the word and all other sensations. The British theoretic and critic Josephine Machon describes the synaesthetics as a performance in which different artistic forms, principles and techniques are brought together in a way to evoke an emotional response by the viewer.i Looking at the history of the live image it will become clear that the DJ/VJ performance is another way to try to get people through Alice's Looking Glass into another world. To fully understand the importance of the intertwining of sound and image in clubs, I want to draw attention to the fact that the DJ/VJ show has its roots in performance art and therefore deserves special attention when describing and analysing conceptions of contemporary performance art. Because the differences among VJ's are enormous nowadays, I will focus on performances that make a physical and psychological connection with the public through the synthesizing of various media like sound, image, smoke, smell, etc. These synaesthetic performances can be seen as the first attempts to create a virtual reality outside the confinements of the CAVE (Cave Automatic Virtual Environment)ii or specially designed suits, in spaces in which the participation of the public is crucial to the success of a performance.

Before looking at contemporary practices I will briefly go into the history of the live image and its connection to sound, followed by the circumstances that lead to the introduction of the VJ.

1. From miracle plays to magic lanterns

The first experiments trying to transmit sensations to a public can be found in the early days of art and theatre history. As soon as the clergy recognized the value of the image that would educate and influence their public, churches were being painted and people were led into a world of wonder. Outside people were treated to plays that could last for days while being performed simultaneously on different stages. The camera obscura was one of the first devices that presented imagery without religious meaning. Although it was Aristotle who first recognized the possibilities of the machine, the camera obscura was not actually developed until the end of the Renaissance. At first it was used as a drawing device to see the world in perspective, but soon it was also used as an instrument for artistic and miraculous events. Special rooms were built to let people see and wonder at the miracles of life.

Also from the official arts like painting people tried to impress the viewer. Renaissance painters used mathematical rules of perspective, through which they created endless spaces. Painters from the Baroque surpassed these rules and created a benevolence and density of imagery and narration, which were nevertheless best visible from one single point of view. The work of the Panorama painters at the end of the 18th century was a breakpoint in art history. For the first time, details in a painting did not aim for one central focus point, but they were rearranged and held together by the whole horizon. The Panorama painting caused for a democratic way of looking in which the viewer decided what their focus was. The goal of the Panorama painters was to create an image that reflected nature and reality as close as possible. The illusion of an active experience of space was used to give the viewer the sense of being in two places at the same time, that of the painting and of the real space. This idea is similar to that of artists building on concepts for a CAVE environment or VJ performance. Another comparison with the CAVE and VJ performances is the narration of multiple stories in one image. This density of narration, which according to

theorist Lev Manovich was already visible among Dutch landscape painters from the 17th century, can also be found in the work of today's VJ. By accumulating and changing the images multiple storylines are created in which the viewer can choose its own path.

When the first boom of the Panorama passed, the Diorama caused for a new up rise. In 1822 the Frenchman Jacques Louis Mandé Daguerre was the first to present the diorama. Large screens were placed in front of an audience of 300 people on which every fifteen minutes the image changed. During the changes of the setting the public marvelled at the use of smoke-, clouding and light effects accompanied by music. The first synaesthetic performance was a fact. A real use of staging of lights and effects came with the invention of the magic lantern. The first lanterns had already been invented in the 17th century. And although they struck the audience, who thought of it as sheer magic, they came to their true development in the next century. This was mainly due to the heavy and impractical size of the apparatus. With technical improvements the lantern became more transportable and therefore their popularity grew. The lantern projected images on smoke or mirrors that gave a feeling of transparency and immateriality. At the turn of the 19th century this lead to a whole series of especially ghostly performances. Through the rest of the century special buildings were built where different performances were held. Accompanied by music, these performances can also be seen as among the first cinematic experiences. To improve the illusion different lenses were developed for the lantern which made it possible to project overlapping images that gave the impression of movement.iii

There were also different approaches with regard to connecting music and light. In 1725 Louis Bernard Castel writes about a 'clavcien oculair'. This ocular harpsichord (later better known as color organ) would consist of two colored discs that were connected to a harpsichord. Castel investigated the analogies between light and sound and with this organ he wanted to paint with sound. Although he never managed to build his organ, various people picked up on his writings. Castel based his findings on observations in the field of physics, but the psychological and philosophical implications form the essence of his research: "Not in dreams, but especially in the state of dizziness preceding sleep, or after listening to music for hours, do I feel the correspondence between colors, sounds and scents. It seems as if they all rise mysteriously form the same ray of light and, subsequently, reunify in an amazing concert. The scent of deep red carnations above all has a magical effect on me".iv He wanted to use the organ as a means to paint with light. Although he never actually managed to build his organ, to this day his writings are still popular with many artists that are looking for ways to (scientifically) make a composition by using light. Descriptions of the sort as described above are also reminiscent of the effects of psychedelic drugs that were in use during the Romantic area and later in the second half of the 20th century.

Another important influence in the history of synaesthetic performance was the German opera composer Richard Wagner. In 1849 Wagner proposed the idea of a Gesamtkunstwerk in his article 'The Art-work of the Future'. Wagner believed that the future of music, music theatre and all the arts lay in embracing the Gesamtkunstwerk, a fusion of the arts. He used his operas as a vehicle to create an immersive experience in which, as Wagner wrote, "the spectator transplants himself upon the stage, by means of all his visual and aural faculties." A few years later he opened his own theatre, the Festspielhaus in Bayreuth, Germany, where he reinvented the conventions of the opera.v

It is not surprising that although the inventions were very influential, in their time they were regarded as outside the scope of Art. They were not regarded as representing reality in a true sense, but were merely deceiving and suggesting the preternatural. Only in Romanticism were they appreciated and seen as works of art in which you could dream and be overwhelmed. In the late Renaissance they had merely been seen as experiments and tools for artists and scientists. When the Realists took over in the early 19th century they emphasized the technical aspect to dismiss the inventions. The technological influence made them unreliable. Works of art had to be true depictions of nature and could not be touched by artificial instruments. In the minds of a lot of people the mechanization would lead to a decline in cultural values. Nonetheless their popularity grew.

2. New means, different ways

The technological progress that started in the 19th century could not be stopped. Eventually the far-reaching possibilities of print, photography, film and radio transmission lead to new ways of thinking. The technological improvements and the more ready access to different and new media influenced the making, perception and distribution of art. New things were possible because new electronic developments were taking place. Different environments were created through new interpretations of light, image, sound and motion. Through technological inventions people could communicate more easily and express and exchange ideas on a broader scale. Old traditions were questioned and society itself came under interrogation. In the early 20th century all the insights of new ideals and developments came to an explosion. At the same time the rise of a mass economy started.

Counteractions showed themselves in the ideas of the Dadaists, the Symbolists, the Futurists and the Surrealists. These groups reacted against the massification of a changed society and especially against the institutionalization of art. Although these groups had different goals and used different means, an overlapping concern was to open the eyes of the public by showing them a different world in which art was merged with daily life. In some cases the audience was urged to take the setting into their own hands, by asking them directly to intervene in the performance or installation. All these efforts were intended to create a sensation and to make contact with the public through different means. Most of the undertakings failed; instead of getting away from the art world, the art world welcomed the new approaches and the efforts to shock were seen as refreshing new ideas.

After the Second World War actions to create a synthesis between performer and public were nevertheless intensified. In the minds of John Cage and Allan Kaprow the public itself had to participate in the performance to get the most intense experience. In destroying theatrical conventions they anticipated the synaesthesia similar to discotheques and the experience of

participating in mass spectacles. These happenings, performances or installations were new forms of art where the participation of the viewer was seen as a necessity. An important aspect of these Happenings was their role outside the art world. The word Happening was chosen by Kaprow to underscore the role of these performances and to avoid any connotation with the art world. According to some new media theorists, it was this transformation in art that prepared the ground for the interactive computer installations that appeared in the 1980s. Moreover I would suggest they also set the ground for DJ/VJ performance that also came into development in the 1980s.

3. The influence of popular culture

With the introduction of television and video artists increasingly became concerned with bringing art and technology together instead of art and life. What were called 'Intermedia' events took the place of Happenings. The Intermedia performance is described as drawing from both theatre and cinema. By walking into the performance environment the audience would set off detailed film images and activate overlapping time projectors, thus changing the imagery and the time frame of the performance. Although this caused for a lot of suspense and excitement and was more than anything a sensory experience, it didn't call for a total illusion where the viewer would be absorbed into a different world. These Intermedia events can be seen as variations on the early dioramas, magic lantern shows and color organs. The performances were still linked to political ideals. Artists tried to wake people up, moving them from their couches and showing them what else could be done with the existing and growing mass culture.

Andy Warhol was one of the first artists to bridge the gap between popular culture and art. In 1966 he created the show 'The Exploding Plastic Inevitable' (EPI) to introduce the music of The Velvet Underground & Nico in the United States. This show can be seen as the first step towards what we now call a DJ/VJ performance. During the whole show there was a synaesthesia of music, image, color and light. The cinematic homage to the EPI by Ronald Nameth managed to capture the experience in which form and content became synonymous. The film was made in the tradition of synaesthetic cinema, a tradition in which one is made aware of the process and effects of perception; that is the phenomenon of experience itself. Synaesthetic films are made out of different layers and fragments in such a way that one image is continually transforming into other images. Images are being orchestrated in such a way that new realities come out of them. This process was later used by many VJ's. In Warhol's own 'club' called the 'Factory' parties were given where performers and public mixed, the one being indistinguishable from the other. These parties and other newly opened discotheques were linked to the psychedelic movement that also started in the 1960s.

The psychedelic movement is mostly know for its experiments with hallucinogens, but more important, it also helped to shape a real 'youth culture'. At the time immense changes were taking place in America's culture and political field. The civil rights battles, the assassination of John F. Kennedy and the rising anger over the war in Vietnam caused a lot of distress. Especially the younger generation was turning against the system. The younger generation was beginning to create for itself a sense of identity and empowerment that was unprecedented. They were trying to make their own lives and rules, different from the conventions of the mainstream culture. Drugs were becoming a means to escape, their hallucinatory effects were creating a new consciousness far away from the conservative society they lived in. Pop and rock music played an important role in the liberation process. The music was largely inspired by mind-expanding drugs like LSD and marijuana. The psychedelic rock movement continued to expand during the '70s, but it lost most of its hard core after the early '70s when rock groups toned their music and lifestyles down. It returned however in the late '80s and '90s when a new sound - techno and electronic music - was performed at what came to be called House Parties.[vi] Once again the music was intensified by the taking of a new drug, MDMA or Ecstasy.

With the arrival of cheaper equipment which facilitated the production of visuals, next to the DJ there was also a VJ showing abstract and surreal visuals that reacted to or fused together with the beat. The origins of the House movement lay in a belief, a belief in the self: "It was a personal liberating experience with a slow, primal beat and rhythm. 'My house is your house and your house is mine.' House culture was family."[vii] Although the House Movement was very similar to the psychedelic movement in the '60s, life itself had changed, influencing the meaning of these new 'raves'. Taken over by the commercial world, the intentions and feeling of the parties changed. As the parties in the '60s were driven by an inner ideology to broaden self-consciousness, the raves in the '90s were a reaction to a deteriorating society. The raves were a place to get rid of the anger and frustrations of everyday life.

4. The shift to music

Going back to my initial argument that the DJ/VJ act is a recent form of synaesthetic performance, I would like to describe a few examples of Dutch DJ/VJ shows that have been very influential in the last years. By exploring these 'new' performances the resemblance with earlier examples will become clear. These performances are not only similar to the ones before, moreover they also seem to follow a similar path. Not long after the first real discotheques opened up in the '70s experiments with light and visuals appeared. People used liquid-slides, disco balls and light projections on smoke to give the audience new sensations. Some of these experiments were linked to the music, but most of the time they functioned as decorations. These first experiments are very similar to the early history of the live image as described above. As the clergy did hundreds of years ago, now again people tried to lure an audience into their world by affording them new sensations. It was not until the late '80s however that an intense relationship between sound and image occurred. This was mainly due to the availability of cheaper equipment like projectors, video cameras and players and the arrival of affordable computer technology.

The fusion of image and music coincided with the arrival of new sounds. 1977 marks a time in history when a new technological tool - the synthesizer - was introduced, and becomes paradigmatic for this shift in attention. In that year bands and singers from different musical backgrounds began making use of

this new technology. It was a German group, Kraftwerk, that bridged the gap between the rock music of the ‘60s and ‘70s and electronic dance music of the ‘80s. With regard to synaesthetic performance the music was more than ever calling the shots from that moment on. The year 1981 marks another important step that was of importance for the intertwining of sound and image and the future of video. Music television station MTV (followed a few years later by others like TMF and the BOX) started non-stop broadcasting of music-video clips. These clips were intended to boost sales on the music charts. In the meantime the clips also provided the singer or band with a more profound image. By overturning traditional conventions regarding imagery a new visual language was created which reached thousands of people at the same time. With the coming of digital video editing this became even more apparent. Regardless of all that can and has been said about the advantages, disadvantages and meaning of this new phenomenon, the fact remains that it did lead to a stronger connection between music and visuals. A number of VJ’s today still say they were inspired by all the visual manipulations and effects in music-video clips.

Around 1985 the House movement and House music arrived in The Netherlands. At first it was only popular with a small group of people and it mostly got played in clubs in the west of Holland. After the second Summer of Love in 1988 things changed and the new Acid House became widely accepted and very popular. The spirit of the House scene was - as elsewhere - one of togetherness, happiness and peace. This was reflected in the staging of the events. The House Parties were large gatherings of people who came to enjoy the music, there was no band or anything to watch, just a pile of electronics with a DJ mixing other people's music. To give the parties a more profound look, a face or even an icon, the VJ came into play and complemented the music with visuals.

5. Pioneers of the Dutch scene

It is hard to determine who was the first VJ in The Netherlands to introduce visuals on the dance floor. If it is primarily about visuals being presented in a new context, the club Mazzo was one of the first places to experiment with imagery. Beginning at Mazzo in 1979, experimental filmmaker and VJ Peter Rubin is credited with being the world's first live-mixing VJ. He developed the art form while trying to relate the new possibilities of technological visualization to young music-oriented cultures in a humanistic, socially oriented manner. It was not until the introduction of House music however that visuals become aesthetically synonymous with the music. Different people started around the same time in different places, but all had the same goal: trying to create a real-time continuity between image and sound. The aesthetics, goals and material they used were as varied as the people producing them. A few examples:

Micha Klein, one of the VJ's who is also known outside of The Netherlands, likes to create a synaesthesia of beauty. Making all the visuals by himself on the computer, Klein wants to create positive icons that are uplifting and will “give you a boost and a positive vibe for the whole day.”viii With all the bouncing colors and whirling forms, his work reflects nostalgia for the psychedelic movement in the ‘60s when everything was about love, peace and happiness. Most of the animations he makes are based on themes which are formulated together with a DJ. The decision about which image to show when is taken during the evening and depends on the music and the mood of the audience. Although Klein is often described as being too superficial and commercial, his work has been of great importance for both VJ performance and digital photography.

Geert Mul and Titus van Eck, initiators of Cut Up and D.U.M.B. (Dutch United Media Base) had different ideas. The two organizations were set up with different goals. Cut Up was the more commercial and D.U.M.B. became a platform for artists from various backgrounds. The members focused on the creation of specific environments to try to bridge the gap between the audience, the music and the visuals. By hanging cameras and microphones in the space and mixing those sounds and images with their own they tried to interact with the public. These experiments recalled the Intermedia events of the ‘60s. But instead of showing people the effects of mass media or the uses of multimedia D.U.M.B. tried to change existing formulas of presentation. They wanted to break out of the club scene to develop the VJ potential into non-traditional ways of presentation. According to D.U.M.B., removing a traditional context and showing different things each time would enhance the experience. "It's all about creating a different sensation".ix The visuals of Cut Up reflected a similar approach. By using existing images from films, documentaries or news items they deconstructed information given by the media, to create a different meaning.

Another VJ who has been very influential is Gerald van der Kaap. Apart from being labelled the Dutch Godfather of VJ, Kaap has also been regarded as the VJ artist who introduced the image in a different way. With his background in art he tried to create a total environment where the music, the space, the light and the visuals intertwined. He is one of the few VJ’s who is not afraid to turn off the screens or to mix his own sounds through the music. First and foremost he is looking for ways to surprise the audience and to shake them up once they seem to be losing their attention.

This showing of new realities was also the starting point for Jaap Drupsteen. Drupsteen’s main interest lay in the visualization of music. He tried to let people see what they were hearing. He would not merely follow the music, but through a multi-layering of video images he orchestrated the visuals thereby allowing the images to materialize in as many layers as the ensemble itself. The composition would change with the rhythm of the music: forms, colors and perspectives changed, sometimes gradually, other times abruptly. This way he revealed music as it is - a comprehensive structure that develops over time. Drupsteen’s work clearly reflects the ideals of a Gesamtkunstwerk as Wagner sought its performance in the 19th century.

6. The second half of the ‘90s

By the early ‘90s House music had become more than just something to dance to. It was the gateway to collective euphoria, a huge shared secret and a massive in-joke in comprehensible to the mainstream. A whole generation was in on it, meeting at

motorway service stations in the middle of the night to follow coded directions to illicit parties and dance until dawn. The police and local government officials hunted down these outbreak of outlaw spirit that spurred hundreds of thousands of people to break into warehouses and set up sound systems in remote fields. But in the mid-'90s by then the more generally termed 'dance music' retreated back into the clubs, opting for constraint and control, and in the process created its first generation gap.

The art of the VJ of the second half of the '90s had become as diverse as the different styles in music. In the second half of the '90s everybody seemed to work with digital technology and everyone could be a DJ or VJ. The emphasis on the space and the context of the presentation, things that were still important to the early VJ's, was replaced by large screens that hung around the dance floor. This stands in sharp contrast to the Happenings and Intermedia events, when visuals were projected onto the people, which caused the audience to connect with each other. The background of the VJ had also changed. The pioneers of the Dutch scene were artists who had been trained in art academies. The second generation of VJ's formed collectives with people from different traditions. Some had their roots in computer programming while others in graphic design, film directing or sound. These collectives were the perfect example of the cross-disciplinary collaborations that found their heyday in the mid-'90s.

Just as all of the dominant music genres (trip-hop, drum and base, big beat, ambient, etc.) depended heavily on technological progress, many VJ's now started editing their material live on a computer which could change and recompose the material in many ways. Jeroen Hofs (Eboman) was one of the first to make full use of the possibilities of the computer. By using samples of rock, house, funk hiphop and jazz and mixing these together to connect them into a new sound, he created the ultimate form of cross-over. He applied the same principle to video. Like the sound, the images were in stereo; the result was a total immersion in the story. He also connected sensors to his arms which would allow him to trigger sounds and images with the movements of his body.

A collective that used found footage and other film material to tell their story was Captain Video. The collective consisted of a 3D animator, editor, graphic designer, director and a multimedia designer. When they just started they primarily used graphics to impress the audience; the faster and more colorful their images were the better. Noticing that this strategy was not very satisfying, they returned to the use of pre-recorded video fragments. Out of this existing material they made new stories, sometimes commenting on the mass media other times to spread a message. Their way of working very much reflected the '90s postmodern culture that was reworking, recombining and analyzing already existing media material to make sense of the world. Coming into the new century Captain Video moved to creating live movies out of existing material that developed together with the music of the DJ., showing new ways of telling a story. It is an approach that can also be found in experimental cinema, video- and computer games and internet.

Pips:lab is a collective that use the computer to combine and connect different disciplines in a new way. One of their main goals is interaction with the public. They are not satisfied with mere projections of visuals that react to the music or vice versa. They work in the tradition of John Cage and Alan Kaprow in a way that the success of the performance depends on the participation of the audience. But whereas during the Happenings the audience was merely seen as material, Pips:lab gives the viewer tools to interact directly with the performance. The player has to perform actions to move the visuals and the narration of the 'play' forward. The public gets to decide in the end what it does or does not want to see. It is not surprising that a new incentive has tried to involve the public. Developments in computer based media have transformed the viewer into an active user. An image became something a user would actively go into, zooming in or clicking on individual parts. The image was no longer static but had become interactive. The performances of Pips:lab remind us of theatre plays where the staging and technique orchestrate the viewer's attention over time. But instead of being spectators, the public have now become participants.

With the beginning of the 21st century a renewed interest in the olfactory surfaced. The VJ collective Barkode wants to create a synaesthetic performance by triggering the subconscious. Their show is a succession of encounters, chance meetings of words, images and sounds. Many transparent layers overlap, leaving the viewer lost in time, space and emotion. But as time passes, more story lines develop, which make it even more unclear: does the spectator find the plots or are the consecutive images and sounds leading up to something? Barkode describes their shows as 'confusing constructivism', inspired by the subconscious. By using smell they augment the atmosphere; pleasant smells heighten the experience, and by employing foul odours a feeling of disgust is triggered. Their performances are a postmodern version of the experiments done in the Romantic area and in Sensurround films of the early 20th century, when smell cards were given to the audience to accompany the sensations in the film.

The potential of the Internet is used by a group of VJ's who call themselves Snowcrash. Through live streaming Internet connections, participating musicians and VJ's world-wide are connected to a live stage. Their play is mixed with the music on the stage and the visual input is projected onto various screens. Special developed electronics assure that everything is synchronized. To still be able to unravel the different cultures and to grasp the global effect they ask the participants to comment on specific topics, ranging from political statements to examples from their popular culture. The musical and visual battle is played out on stage and can also be followed live on the internet in the confines of the living room. Snowcrash exemplifies how the VJ experience of today is not just a fusing of different disciplines, but is also an expansion of national borders.

7. Moving into the world

Up till now I have mainly concentrated on the Dutch VJ scene. This seems a logical step as Club Mazzo in Amsterdam in 1978 was one of the first clubs in the world that facilitate for and believed in the impact of using visuals during their nightly dance parties. The problem most clubs had was hardware. Existing

hardware was too expensive to buy and was only used by advertising agencies, national broadcasters and government agencies. After developing their own video mixing tool, Mazzo embarked on a unique cultural path in the history of clubbing. Their goal was to develop new theories and practices regarding visuals, music and social ideology (i.e. how to best communicate social messages within the rapidly changing technological environment). At the heart of these experiments was the presumption that the power and scope of sound and image in perfect balance could best meet the needs of these latest challenges. One of the primary goals of the Mazzo experiment was to establish an environment wherein one would not run away from reality but, rather, get the inspiration and renewed mindset to improve the conditions that exist within reality. The underlying concept rested on fact that the video mix was not conceived as an extension from the world of music or art but, rather, was developed as a form of progressive social communication. However, as culture has become more and more commercialised, the social messaging which permeated the initial period of video mixing has been replaced in great part by the flashings of the individual VJ.

As I said before the art of VJ has gone into many different directions, both content wise as well as form but the experience also crossed many borders. Every club in the world has its own VJ or team of VJ's, it will go too far to mention the differences or similarities between the various countries or people. What is striking however is the influence of the technological advances. These developments have offered every country of the world new and profound possibilities for intercommunication and interactivity. As a result of this progress, an absolutely vast number of dance clubs have sprung up in every conceivable corner of our planet. In one form or another (primarily due to the Internet), every one of these countless number of clubs is also directly connected to another, meaning that there is, in fact, an established connection between every club in the world. Such a structure is unprecedented in history. To an extraordinary degree, the predominant social philosophy within the youth of these clubs transcends the individual socio/political directions of their respective governments, allowing for the possibility of an enormous and unified global voice for social reform.

Striking in this respect is the influence of the so-called 2nd and 3rd world countries over the 1st world countries. This is not so much by means of new visual languages or graphics that are being shown, but more importantly a new vibe seems to surface. Coming form a very different background and cultural status the attitude is very different from the established 1st world VJ's. This also effects the audience watching the performance. The resulting performances carry a very special, underlying electricity that is simply not present in 1st world performances. An experience that is totally invisible. But totally visible. Hopefully this will provide a significant, positive influence and renewed inspiration.

8. The future of visuals and sound

By going through history I have tried to shed light onto the phenomenon of VJ culture. Every now and then I also briefly hinted to the influence of VJ on other disciplines, but foremost I tried to show the history or the origins of the synaesthetic performance. When looking at recent trends in VJ it would not be an understatement to say that VJ indeed is a new form of synaesthetic performance, a form that might develop in interesting ways. At times it may be thought that history is merely repeating itself, the only difference being a new medium. But as I have pointed out, one thing has changed dramatically and calls for new research and interest in this area: the audience is not composed of a few insiders anymore, but the whole synaesthetic experience is massified. Although there is no telling where the VJ performance will lead us, already questions arise which call for further investigation: What is the influence of all this media circuitry on existing disciplines, on the position of the performer and of the audience? What will happen with the influence of performers from non-Western countries?

When looking at the speed of technological implementation, activation in the development of a new, positive, contemporary consciousness seems necessary if we ever want to step out of an ever-expanding Looking Glass. One important area to focus on is the club scene, as it has always been the environment where youth culture gathers together most consistently and in its most significant numbers.

Thanks to: Peter Rubin, Amsterdam

More information on VJ http://www.visualsensations.nl

References

[1] Gregory Battcock and Robert Nickas (eds.), The Art of Performance. A Critical Anthology (New York, E.P.Dutton, Inc. 1984)

[2] Jim Cook, 'Shallow happiness never lasts', in: Flash Art, October 1999

[3] Doug Hall and Sally Jo Fifer (eds.), Illuminating Video. An Essential Guide to Video Art (New York, Aperature Foundation, 1990)

[4] Lynn Hershman, Clicking In. Hot links to a Digital Culture (Seattle, Bay Press, 1996)

[5] Localizer 1.0 The Techno House Book (Berlin, die Gestalten Verlag, 1995)

[6] Localizer 1.3 Icons (Berlin, die Gestalten Verlag, 1998)

[7] Josephine Machon, (Syn)aesthetics and Disturbance - A Preliminary Overview, in: Body Space & Technology Journal (Brunel University, vol.1, no.2, 2001)

[8] Laurent Mannoni, The Great Art of Light and Shadow. Archaeology of the Cinema (Exeter, University of Exeter Press, 2000)

[9] Lev Manovich, The Language of New Media (Boston, MIT Press, 2001)

[10] Jacqueline Oskamp, 'Veejays. Beeldende kunst of behang?', in: THD 02, zomer 1997

[11] Helen Paris, (Re)Confirming the Conventions - An Ontology of the Olfactory, in: Body Space & Technology Journal (Brunel University, vol.1, no.2, 2001)

[12] Peter Shapiro (ed.), Modulations A History of Electronic Music: Throbbing Words on Sound (New York, D.A.P., 2000)

[13] Gene Youngblood, Expanded Cinema (New York, E.P.Dutton & Co., Inc., 1970)

[14] Andrea Zapp and Martin Rieser (eds.), New Screen Media. Cinema/Art/Narrative (London, British Film Industry Publishing, 2002)

---

[i] Josephine Machon, '(Syn)aesthetics and Disturbance – A Preliminary Overview', in: Body Space & Technology Journal (Brunel University, vol.1, no.2, 2001)

[ii] The CAVE creates a three-dimensional virtual reality that allows users to walk around in whatever environment they choose. Images are projected onto three walls and the floor, completely surrounding the user's visual field. With special 3-D glasses that track where the user is looking, it is possible to actually walk around and behind images.

[iii] Although the magic lantern has always been regarded as the precursor of the cinematic experience, it has even more similarities with VJ culture in our time. This is especially true for the early experiments where the magical and illusionary effects were very important.

iv Frans Evers, "Transcript of the Permanent Flux lecture", in: Permanent Flux (De Balie, Amsterdam, 1999)

v http://www.artmuseum.net/w2vr/contents.html

vi Even though the name 'House' was originally adopted from the legendary 'warehouse' in Chicago, the 'Spirit of House' developed in the '80s due to the more frequent use of private houses for parties. This situation arose because either they could not get past the bouncers in established clubs, and/or these clubs were too expensive. Source: Localizer 1.0 The Techno House Book (Gestalten Verlag, Berlin, 1995)

vii Peter Rubin, "Chromapark and Beyond", in: Localizer 1.0

viii Jim Cook, "Shallow happiness never lasts", in: Flash Art, October 1999

ix Jacqueline Oskamp, "Veejays. Beeldende kunst of behang?", in: THD 02, summer 1997

Messages For A First Person Perspective

Maia Engeli

Planetary Collegium, University of Plymouth, UK
Architecture, Technical University of Delft, NL
Hyperwerk, University of Applied Sciences, Basel, CH

maia@enge.li

ABSTRACT

This paper is about explorations into the creation of powerful messages within the first-person shooter game Unreal Tournament. Unreal Tournament is interesting because of its popularity, large active community, affordability, and largely open-source accessibility. Creating messages in this environment puts the author in a twofold situation; with the interesting possibilities of VR editing on the one side, and on the other the challenge of having to create a message in an environment filled with predetermined expectations towards an ego-shooter game.

Keywords

First-person shooter games, computer games, game editing, interactivity, visual expressiveness

1. MESSAGES IN FIRST-PERSON SHOOTER GAMES

Within first-person shooter games very powerful "messages" can be created. Examples later in this paper will illustrate this. These are the results of workshops with Interactive Media students at the Hyperwerk Basel, Switzerland and Architecture Students at the TU Delft, Netherlands.

The messages are not communicated through words or iconographic images, but through an expressiveness composed from visual, aural, and interactive aspects in a dynamic digital space. The messages are experienced through interaction. Since there is no clear, culturally based language with a long tradition, like spoken language, the resulting messages are more ambiguous. But they are understood thanks to the increasing "maturity" of the player, who has progressed from a reader of texts or a viewer of images to a user and now a player [1]. Or as De Kerckhove describes it "This is the passage from "point of view" to "point of being": the "point of view" expects for each body to observe the world from its interior; the "point of being" implies an intense exchange between a single individual vision and the multiplicity of exterior visions." [2]

2. POPULARITY OF FIRST PERSON SHOOTERS

First person shooter games are nowadays the most widespread VR application and many of the game engines, like Unreal or Quake, offer easy to use editors for level design. This makes them a prime media for popular VR entertainment and creative use of VR. The community around the different ego-shooter games is large and multifaceted. It operates globally through the Internet and includes many levels of communication; playing games together, exchanging knowledge and data, as well as gossiping.

The aspect of popularity makes first-person shooters an interesting media to work with, because of the large audience that can be reached. Messages can be created in artistic ways, expanding existing realms of knowledge and experience, or used as bluntly as in America's Army's free first person shooter game that invites you to "Become a member of the world's premier land force ..." [3].

3. SPECIFIC QUALITIES

Ego-shooter game levels have a particular aesthetic; they often have a "medieval" look because of the rough stone textures and sparse, low-energy light sources that are generally applied to the environments. The aesthetic possibilities are restricted to allow for fast graphics, but there are still a lot of possibilities regarding form, textures, illumination, and behaviour.

For architecture and interactive media students the design of game levels introduces a number of interesting new features and possibilities to the design process: 1) The user/client is the PLAYER. The player wants entertainment and fun. The player needs to be able to understand the possible actions. Player motivation becomes a key concern for the designer. The design of an ego shooter level has to focus on the creation of an interactive experience and the dialog between player and environment. 2) Levels are BUILT – 1:1 – real, not virtual – they are more than plans or mock-ups of a design – levels are real game environments for real players. 3) Levels can be MODIFIED. Modifications in response to comments from players become part of the design process. The quality of the final design increases through the enhancements that are introduced in response to the players' feedback. 4) Levels are DYNAMIC and REACTIVE. They may look like stone, but they are made of bits. They can behave, trigger the player's actions, and respond in various ways.

4. SEVEN SAMPLE MESSAGES

The following examples are the results of two workshops, in which the editor Unrealed2 and Unrealed3 were used to create levels. Special emphasis was put on exploring the specific expressive power provided by the first person shooter game.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

4.1 Workshop: Hyperlevel

http://www.hyperwerk.ch/~hyperdat/unreal/

This workshop was about exploring the expressive power, without a predefined theme for the messages. The first two examples elaborated on known aspects of ego shooter games and their 'reframement', so that the usual behaviour within a first-person environment becomes either destructive or especially pleasant. The third example focuses on special VR qualities, like mirrors, changing environments, and hyperlinks between different spaces, to translate a statement into an experience.

4.1.1 Drift or Die

This is a counterintuitive level, which asks the player to just let herself drift and enjoy the ever-repeating journey through a meditative sequence of rooms. Self-initiative is punished with one of the different ways to die that are modelled into the level, i.e. being squeezed by a platform or falling too far down.

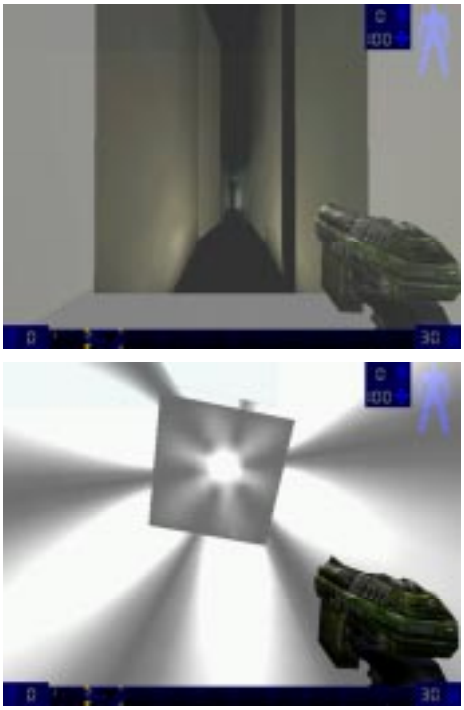


Figure 1: "Drift or Die" by Irena Kulka and Franco Schwörer

4.1.2 Dream Day

This level is built as an absurd, kitschy, two-story house filled with cliché loaded images providing a dream world that can be leisurely explored. The atmosphere is just too pleasant so that the playing of a deathmatch within this environment actually amounts to a welcome contrast.

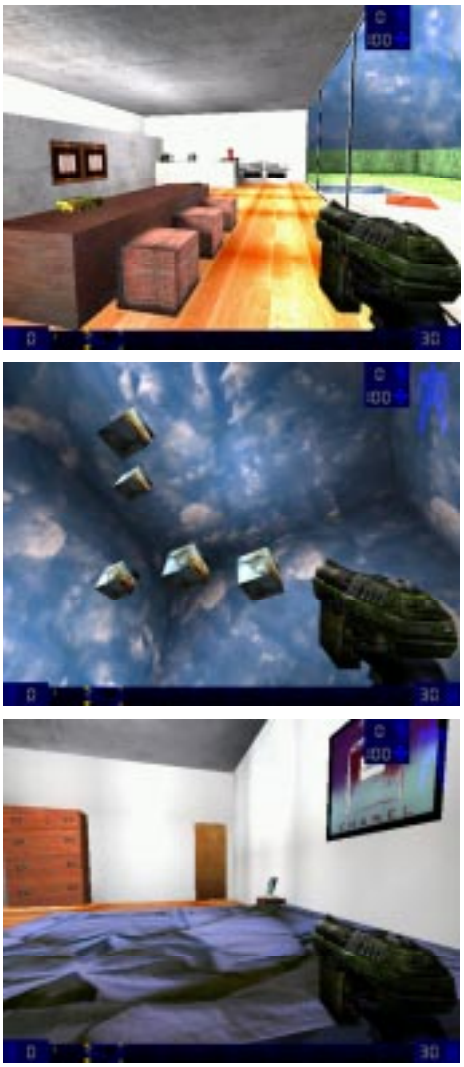


Figure 2: "Dream Day" by Marc Dietrich and Michael Huber

4.1.3 Darwin is Dead

This level is about emphasising Chaos Theory over Darwin's theory. The textures and the use of mirrors make it hard to recognize the shape of the space. At some specially marked locations one gets teleported into a very different space, patterned with fractal images. On return one of the "Darwins" on the centre volume is replaced by an image from chaos theory.

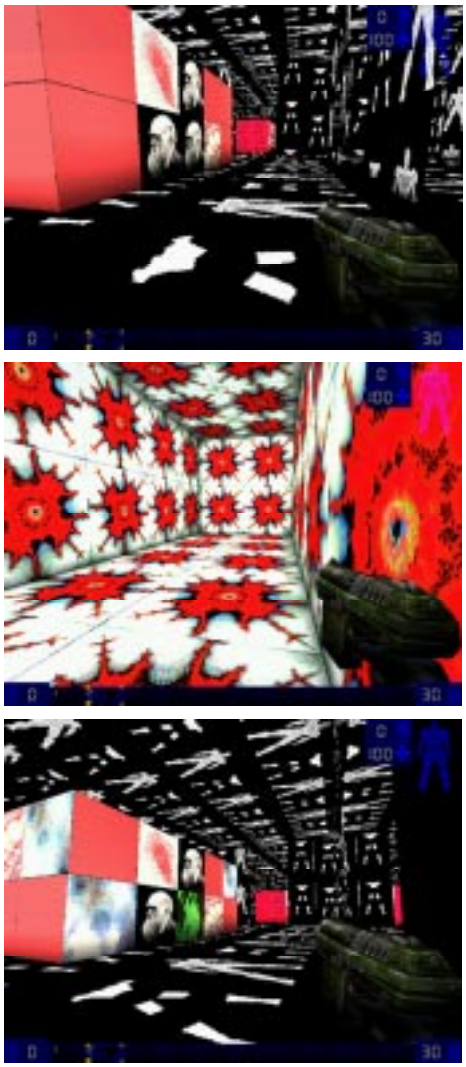


Figure 3: "Darwin is Dead" by Anja Kaufmann and Julia Kehl

4.2 Workshop: Mediated Discourse

http://maia.enge.li/gamezone/parkstad/

The focus of the Mediated Discourse workshop was to create messages about Parkstad, an urban project in Rotterdam. The main characteristics of the project are its location between different ethnic communities, the high density of different kinds of traffic in the area, and the aim to create a high-density district with a 'large' central park.

4.2.1 Under Construction – Creating your own Parkstad

In this level at first only the layout of Parkstad is visible, with the potential to make the planned city apparent. The buildings will grow if you step on them, which will lift you into the air, offering a new perspective and a "parachute" flight back down. The appearance of trees can be triggered as well.

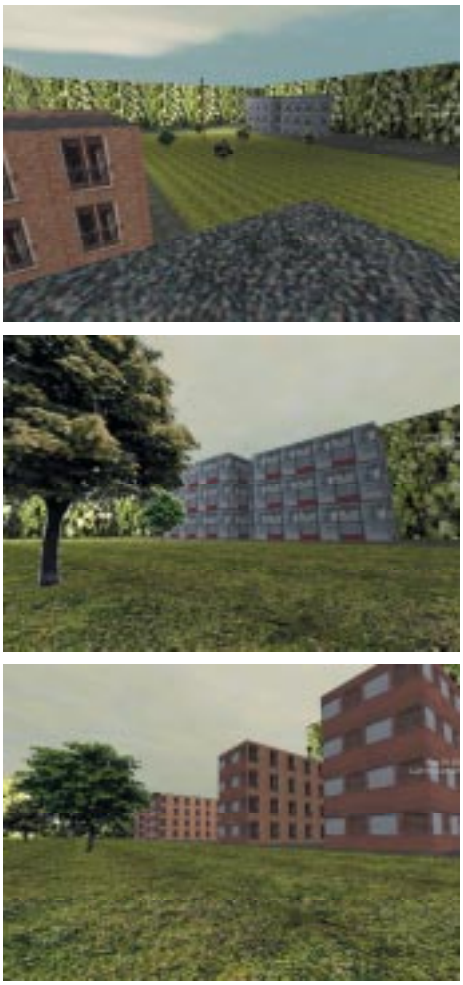


Figure 4: "Under Construction–Creating your own Parkstad" by Rob Kotte

4.2.2 Hidden Islands – Expanding Perspectives

In this level the city does not show its parks, but the birds give you hints about where the green islands are hidden. Each of the five parks represents a typical "park" theme, like "meadow with trees", "zoo with rabbits", "an exercise area for a sports team", "a playground", and "a pond". All living beings in this level can be shot, the consequence of which is the spreading of nervousness, items out of place, and chaos in the whole system.

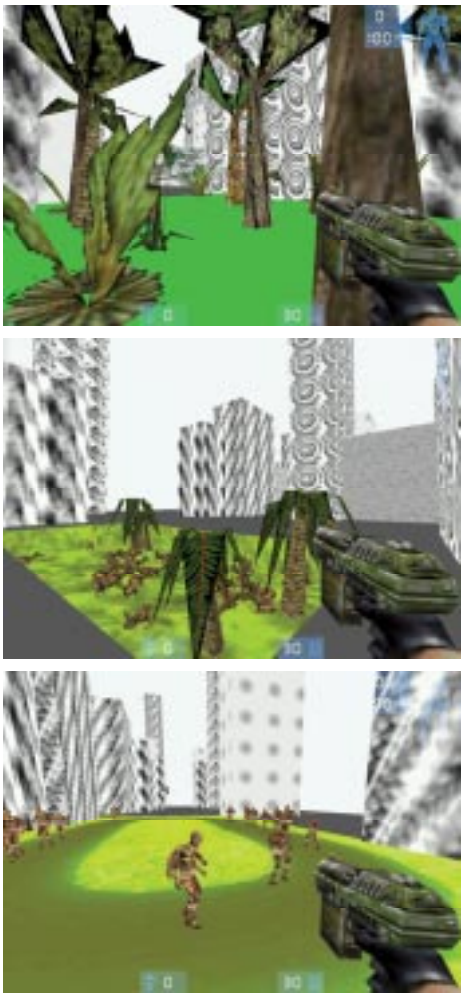


Figure 5: "Hidden Islands–Expanding Perspectives"

4.2.3 Road to Parkstad – An Endless Drift

The multi-lane road within the planned Parkstad in Rotterdam will cut through the area very dominantly. This level confronts the experience from the roadside, with the drift along the road through the urban area.

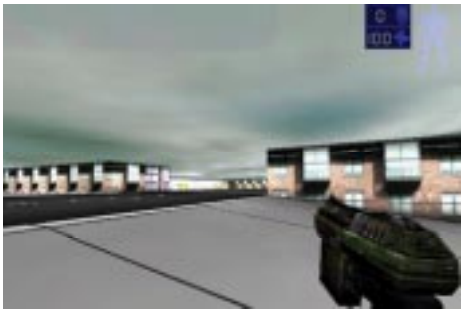


Figure 6: "Road to Parkstad–An Endless Drift" by Bas Plasschaert

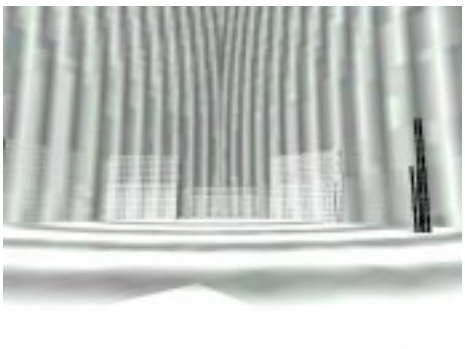
4.3 Workshop: Metaworx Unreal

http://www.hyperwerk.ch/~hyperdat/unreal2/

Metaworx is a planned travelling exhibition showing student work from four Swiss Universities. In the workshop the students were asked to either create a piece to be exposed or to create level that could serve as an interface to the exhibition or parts of it.

4.3.1 Golden Calf: Metareal

This level shows an immersive, interactive exhibition of thesis work. In addition to the expandable exhibition space, there are spaces that illustrate the often painful process of the thesis work. There is the "nimbus" where the player floats comfortably but without control of direction and slowly gets desperate to find an exit. There is "hell"; after the players falls into a dark hole and sees nothing, feels very lost in the dark until he or she recognizes a dim light far away. And there is "heaven" with windows down into the exhibition. The windows show different perspectives onto the works. Visually the level is abstract, mostly black and white, and deliberately aiming at a contemporary graphical aesthetic.



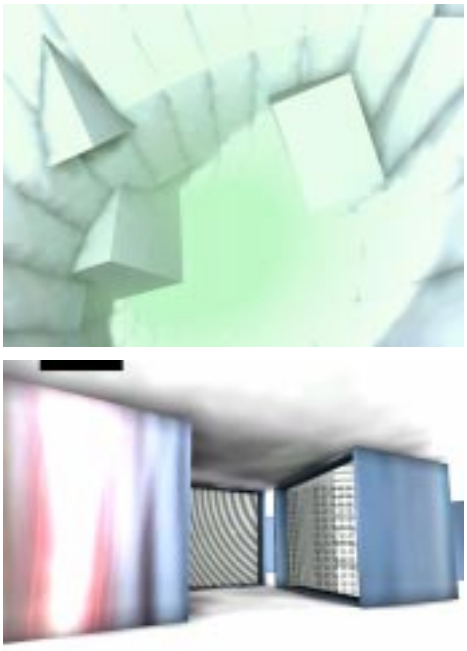


Figure 7: "Golden Calf: Metareal" by Luca Vincente

5. INSPIRATIONS FROM THE ARTS

As with any new media it is the artistic approaches that serve as eye-openers by indicating further potential power of a media. The following three examples where important inspirations for creating a highly exploratory atmosphere in the workshops.

1) "NextLevel" (by Miriam Zehnder, Eric van der Mark, Patrick Sibenaler, ETH Zurich, http://caad.arch.ethz.ch/~patrick/LOCAL/research/playground/) investigates the relationship between game, architecture and narratives. It is a reinterpretation of the novel "Through the Looking Glass" by Lewis Carroll. Alice, the story's main character, imagines different worlds for the different characters of the chess game. The design of the game level was not done as a direct translation of the novel, but "aiming at understanding the essence of the idea and then creating meaningful interpretations" [4]. The environment is hybrid and hyper-linked; it contains contrasts on the visual as well the behavioural level. This makes the journey through Alice's world a real adventure with unpredictable encounters, like: Gravity does not always function as expected; speed of movement is relative; or what seams to be an enemy may turn out to be one's own mirror image.

2) Jodi's "untitled game" (http://www.untitled-game.org/), game levels in which the interface has been reduced beyond the minimum is a "subversion of the aesthetic expectation of opaque source codes and understandable output" [5]. The reduction of the visual aspect to black and white 'pixels' provokes a distorted perception of the movement in the game. Nonetheless, the changing visual patterns and the sounds remain the only level that can be interpreted and after some time one can recognize patterns familiar from the known shooter game. The effect is a heightened awareness of these shooting patterns as well as the insight, that even this most abstract representation can be as terrifying as the realistically rendered game.

3) Max Moswitzer and Margarete Jahrmann dig deep into the possibilities to alter the game and combine it with other information sources. Their newest example "Nybble Engine" (http://www.climax.at) is about hacking the server part of Unreal Tournament and, for example, replace shooting by sending anti-war emails to the president of the United States. They have overcome the discrepancy between shooter game and message and are exploiting the game with all its qualities for their aims. "An aesthetic message is usually the deconstruction of a conventionalized text form or a media text. It is receded by destroying semantic portions in order to increase the aesthetic information." [6].

The arts examples and the workshop examples are quite distinct regarding the depth of exploitation of Unreal Tournament's power and potential. While the arts examples really aim at expanding the boundaries, the workshop results were focussed on detecting the potential expressiveness.

6. OBSERVATIONS AND REFLECTIONS

The results from the TU Delft workshop where exhibited for two weeks at the Architecture School and we realized that we are dealing with two kinds of semantics; there is a very big discrepancy between the understanding of a level by an Unreal Tournament player and the understanding within a thematic discourse, which must originate from the difference in motivation. The player works with a shooting-and-survival strategy, while the thematic discourse focuses on experiencing the message through VR means. To close the gap between the two aims, the discourse level has to be empowered by creating messages that merge with or overpower the shooting-and-survival instinct of the UT players.

From the perspective of semantics the Unreal Tournament player's approach is to read the shape of the space and connect it with possible game strategies, much like a chess player. Moving objects and creatures will draw attention, demand caution, and ask for immediate recognition as enemy, harmless thing, or team mate. Further signs and indicators are either iconic, i.e. health packs, ammunition or weapons, or built analogous to physical occurrences, like doors, bridges, and pathways.

In our message-oriented approaches, we tried to push the semantics beyond the obviously known from the first person games. By introducing novelties, it was also important to consider the need of the player to learn about the novel things and then create the respective opportunity. Few signs or configurations, like the teleporter in figure 8, turned out to be immediately readable.

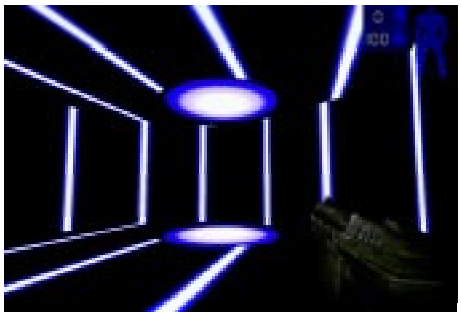


Figure 8: Teleporter in "Nextlevel" (Patrick Sibenaler)

Our experience was that VR semantics, beyond the ones directly translated from a physical reality, are only slowly developing. Players are already very familiar with different modes of navigation also in environments with distorted gravity and different 'speed-zones'. But moving through walls, which is a natural thing to do, from a VR point-of-view, still feels very unnatural to most players.

The messages created in the workshops are less composed from readable signs but aimed at creating situations that can be experienced and understood in connection with past experiences.

In a currently running workshop emotions are emphasized to explore the expression of feelings and search for respective media-specific expressivity.

7. REFERENCES

[1] Engeli M., from within. ***in arquitecturanimació***, Massad, F., Yeste, A. G. Eds., Barcelona, 2002.

[2] de Kerckhove, D., "Searching for the Principles of Web Architecture", in The Charter of Zurich, Barzon, F. (Ed.) Birkhauser - Publishers for Architecture, Basel, Boston, Berlin, 2003.pp. 38-67.

[3] Americas Army, http://www.americasarmy.com

[4] Zehnder, M. Nextlevel, in bits and spaces - Architecture and Computing for Physical, Virtual, Hybrid Realms - 33 Projects by Architecture and CAAD, ETH Zurich, Engeli, M. Ed., Birkhauser, Basel, Boston, Berlin, 2001.

[5] Cramer, F., "jodi.net", in install.exe/Jodi, Baumgärtel, T. and plugin (Eds.), Christoph Merian Verlag. Basel, 2002.

[6] Jahrmann, M. and Moswitzer, M., 2003. Nybble-Engine, Vienna.

850 hPa 01 MEZ

Ivo Flammer
Artist, www.ivoflammer.ch
3, rue Thimonnier
75009 Paris
++31 (0)6 86 83 94 28
mail@ivoflammer.ch

ABSTRACT

Contemporary art lacks a uniform decoding system. On one hand, this lack leads to a formidable creative liberty for the artist; on the other hand, it leaves many viewers puzzled and suspended in search of sensible decoding. In my artwork, the artistic language is contrasted with other languages, particularly that of science, which possesses a precise, well-defined decoding system. Using a scientific object as starting point in my work, I create an artistic interpretation of the object. My interest lies less in the proposed alternative interpretation as such but in raising the viewers' awareness of their own decoding process.

Keywords

Contemporary art, decoding.

1. INTRODUCTION

Contemplating works of art is generally considered to be an intuitive and spontaneous act. However, enjoyment of contemporary artworks is strongly enhanced when one possesses background information on the artist and his concerns. Such supplementary information guides the viewer in his reception of the artwork and facilitates the communication between the viewer and the artist.

Generally, I think the reception and interpretation of an artwork consists of the viewer's projection of the artwork onto his personal decoding scheme. The viewer's decoding scheme depends on previously acquired experience and knowledge: his cultural, educational, and social background, but also the extent to which the viewer has information on the artwork as well as the artist and his works.

As an artist, one of my interests is the language for decoding contemporary art. In my artwork I oppose the language of contemporary art to other languages, particularly the language of science. The scientific language is precisely defined, its interpretation scheme is supposed to be objective and neutral. In contrast, the language of contemporary art is multi-dimensional, emotional and personal. The two languages are combined in my artistic creation through an interpretation of scientific schemes and objects as artistic subjects. My concern is not so much in proposing an alternative interpretation, but more in raising awareness of the viewers' own decoding system which is subject to their background knowledge, experience and personal sensibilities.

In the following sections of this text, I will present two of my artworks, a drawing and an installation. The first one, "850 hPa 01 MEZ 5.1.1998" is an artistic interpretation of a scientific scheme, a weather chart. The second one, "LAIKA", is a painting showing the first dog in orbit. The painting is accompanied by two supplementary pieces of information concerning Laika, a text and a recording.

2. "850 hPa 01 MEZ 5.1.1998"

2.1 Materials

Figure 1 shows a picture of the drawing "850 hPa 01 MEZ 5.1.1998". The drawing is painted with asphalt on paper, its size is 175x150cm. The drawing is an interpretation of a weather chart on which pressure and wind distributions are shown. The specific pressure and wind distribution were taken on January 5th 1998 over continental Europe.

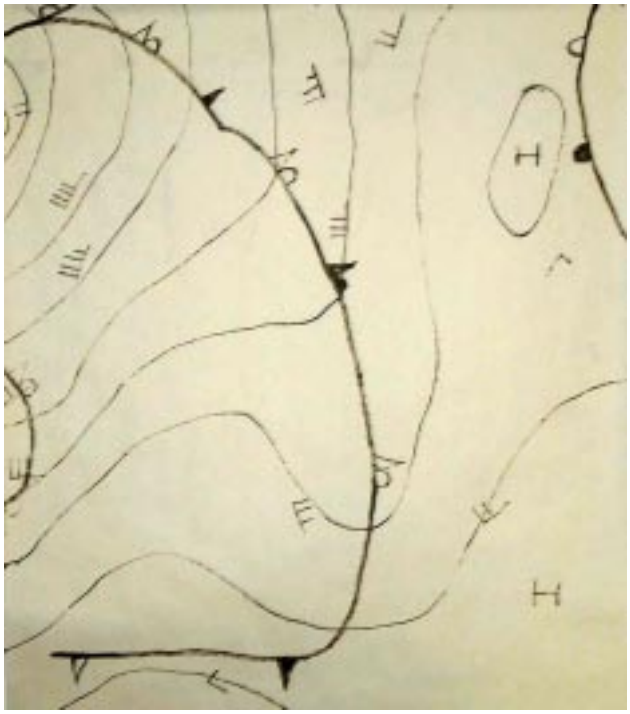


Figure 1. 850 hPa 01 MEZ 5.1.1998

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

2.2 Discussion

The main concern here is the impact of the viewer's level of scientific knowledge on the way he decodes the artwork. Viewers who know metrology will decode the drawing differently from those who have little or no relevant knowledge:

a) A person not initiated to science will not recognize the representation of the weather chart. Such a viewer may read the drawing as an abstract painting of lines and forms.

b) A person initiated to science (which is the typical viewer of contemporary art), recognizes the depiction of a weather chart. This viewer will be confronted by both, the artistic and the scientific aspects of the drawing.

c) A specialist in meteorology will understand all the details of the weather chart (he probably would even find some subtle scientific impossibilities depicted on the drawing). He will imagine the interplay between wind and pressure seen on the drawing. For him the depicted storm might be boring; or he might get the shivers from the visualization of that terrible storm devastating Europe (the depicted storm was actually named "Desirée" by meteorologists).

3. "LAIKA"

3.1 Materials

The installation is shown in Figure 2. It consists of three parts: The main part is a painting made with high gloss paint on plastic. The second part is a text on photographic paper, imitating the font and allure of old, fluorescent computer screens. The third part is a headphone, made available to the viewer. Using a push button, the viewer can start and stop the recording.

The painting depicts the dog Laika lying in its capsule before being sent into orbit on 1957. Laika was the first dog in space; the mission was successful and represents a historical step in the human conquest of the moon.

The text below the painting reads: "On November 3rd 1957 Sputnik 2 was launched into orbit, 500km above the earth's surface. On board was the dog Laika, first animal in space. During her flight in the 105cm diameter capsule, her heartbeat was constantly recorded and sent back to earth. After a week in orbit the capsule overheated and the mongrel dog died.
Laika grew up in the streets of Moscow and was chosen for this mission because of her particularly calm character."

The headphones emit Laika's original heartbeat while she was orbiting around earth. During the mission, Laika's heartbeat was broadcasted live over the radio.

Figure 2. Installation for LAIKA

3.2 Discussion

A viewer approaching the installation first sees the painting on which he may recognize the dog's face. Some people actually do not recognize the dog's face at first glance. Then the viewer starts assimilating the supplementary information belonging to the installation: the painting's tag (title, material and the artist's name), the text below the painting and the recording. The various pieces of information are different in character. The text is written in descriptive language, void of emotion. By contrast, the recording is strongly emotional because it not only transmits the dog's heartbeat but also mentally "isolates" the viewer as he puts on the earphones. As the viewer explores the installation and assimilates the different pieces of information about Laika, his decoding of the painting undergoes constant changes.

Both of the presented artworks - the weather chart and Laika - are concerned with the effects of information on decoding. While the former points to the different decodings resulting from a viewer's level of information on a specific subject matter (i.e. metrology), the latter shows the changing decoding due to the viewer's gradual incorporation of new pieces of information.

4. CONCLUSIONS

Decoding of contemporary art is a complex, multidimensional and individual process. There is no unique decoding scheme for artworks, and this is certainly one of the attractions about art.

Many branches of contemporary art are done by artists with an interest in a specific topic. These artworks may not be fully decoded by viewers lacking knowledge about that specific topic.

New Media introduces a multitude of new technology and new references to contemporary art. With these novel dimensions introduced, New Media art comes to have a specific decoding system. Therefore, viewers not familiar with the technology and references of New Media art are barred from fully enjoying certain artworks.

5. ACKNOWLEDGMENTS

I would like to thank Yoshie Kaga, Markus Birchmeier and Wolf Ka for helpful discussion in writing this text.

melitzah: An Utterance

Risa Horowitz
artist, curator: ace**art**inc.
32-400 Assiniboine Avenue
Winnipeg MB R3C 0Y1 Canada
(204) 478-1416
horowitz@mts.net

ABSTRACT

This paper describes the production of my artwork, **melitzah**, contextualizing the work within semiotics and the use of computers in creating interactive art. This paper has been made for inclusion in the publication of the proceedings of the COSIGN 2003 3rd Annual Conference on Computational Semiotics for Gaming and New Media.

Keywords

art, interactive art, conceptual art, text-based art, melitzah, semiotics, language, interpersonal communication, interpretation, comprehension, lexicography, dictionary, waveform, utterance.

1. INTRODUCTION

I just finished watching the Canadian Broadcasting Corporation's documentary on the World Scrabble Championships, and am completely impressed by the specific skills Scrabble experts have. These experts aren't linguists who are engaged with the use of language as such. Rather, they are obsessive memorizers, pattern recognizers, and are often mathematicians or computer programmers.

The reason I never became a Scrabble champ, even though it is my favourite game, is because I simply don't have those skills. My habit of using the computer to create and manage systems of storage, organization and recall is altogether different. This is a form of dependence, not self-sufficiency, in the same way that my dyslexic friend relies more on his memory than his notebook. Ordinary use of words may fall somewhere in the middle, perhaps: most folks depend on a combination of their memories and on the dictionary when memory or experience fails.

My interest in words stems from a very personal relationship with words and communication. I have always loved words: my earliest memories are of breakfasts with my father, who took great pleasure in hearing me read the daily paper to him. On the other hand, I have always experienced frustration with words: I have glitches, I talk too much, I mix-up words, I lose or cannot find them, and when this happens I find myself enacting a sort of aphasia, describing the function or appearance of the object to which the lost word refers. Example: can you turn the volume up on this, umm, voice amplification machine?

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

We are familiar with the impossibility of perfect communication, though it is the fantasy of becoming the most efficient communicator I can that has led me to create **melitzah**.

2. WHAT IS MELITZAH

My art practice seems to be grounded in worry, repetition, compulsion and endurance. My most recent and current projects revolve around a systematic approach to materials and my desire to collect, amass, and archive.

melitzah is defined by the process of creating it. The work deals with the polysemy of language, interpersonal communication, and personal discipline. **melitzah** hovers between a set of personal questions about legibility, interpretation, expression and understanding, and a semi-lexicographic archival enterprise.

melitzah [the Hebrew word meaning utterance: an uninterrupted chain of spoken or written words not necessarily corresponding to a single or complete grammatical unit][1] is comprised of an audio recording of my voice reading the Canadian Oxford Dictionary, and a visual archive containing the waveforms for the words of the dictionary.

The work takes on three forms:

1. The 22.5 hour-long voice recording of the 65,835 chosen headwords of the dictionary.
2. The 138 volumes of books representing the visual translation of the dictionary into amplitude waveform.
3. The searchable website: viewers may type in words, calling up the waveforms for those words and hearing my voice speak the words to them.[I]

I am curious about these waveforms as visual representations of the auditory, which are linguistic representations of the cognitive. **melitzah** presents an empirically more accurate visual representation of the English language than the letters of our alphabet, which is nonetheless illegible, making analogous the failures of everyday communication with the abstraction of language into waveform.

Semiologically, the waveforms representing the sound of my voice reading the words of the dictionary transcend the arbitrariness of the language and meaning signifiers of our

[I] http://www.risahorowitz.com/melitzah/index.html

alphabet and our combinations of these pictures to make words, to create meaning. The waveforms are mathematically, scientifically grounded, representing the measurable quantities of amplitude, frequency, duration, and volume.

This grounding in the empirical is the basis for my hope-filled attempt to become a more articulate communicator of meaning, a faulty logic lending itself to failure, since the waveforms are ultimately illegible. These waveforms are specific only to me, to my language, voice, inflection, modulation, accent, and cultural and regional influences. Though there would be some mathematical and visual similarities, this vocabulary would change for each and every user.

I came to make **melitzah** as a result of having participated in an experimental audio art compilation, organized by the Canadian media centre Video Verite, and curated by the audio artist Steve Heimbecker. My project involved the digital recording and editing of a list of 'what-if's' I had compiled over the year. I fell in love with the waveforms used by the audio software SoundEdit 16 as the visual interface for the computer user/audio editor. I decided to create a visual vocabulary of waveforms as my next project and set out to determine the source vocabulary.

Although I'd spent many years on an autobiographical archiving project where I photographed myself over a seven-year period[II] , I ruled out my personal journals - by then, I'd had enough of myself.

The story of the Tower of Babel is a myth that has always had resonance for me. This biblical story tells of a strong and united people who, in defense against enemies, endeavor to build "a tower with its top in the heavens."[4] I interpreted this as a figure of speech, but their deity felt this inappropriate since the heavens were his sole domain. He punished the people by creating different languages, confounding their ability to understand one another. I've always felt this was an unfair punishment - after all, how often do you find a strong and united people?

This story was to the point, but too short given my attraction for large-scale projects.

Over the months I began to understand more clearly what was so attractive to me about these waveforms - that they were actual representations of my voice - and I finally decided that the only vocabulary that would suffice was an entire one.

The dictionary - any dictionary - always presents itself as an authoritative tome. Like many archives it is an effective normalizer: the dictionary can normalize the use of language as much as document it. Ideologically, however, the dictionary has two faces: on the one side, it does not question the form of social contract that conventionalizes language use; on the other side, it serves as a resource of discovery for the use of language towards excellence, in providing definitive and pronunciation clues, points of usage and etymology.

The Canadian Oxford Dictionary [1] fulfilled my requirement for quantity - **melitzah** ranges from 2 - 8 gigabytes depending on the audio compression employed, has over 65,000 words, and an audio recording over 22 hours. This dictionary met my desire to engage in a highly structured and systematized studio practice: my work has been influenced by conceptual and process oriented artworks from the Sixties on. It met my interests in language and communication, and it met a self-imposed criteria for Canadian content. I also thought that it would be incredible to create a visual translation of the English language.

3. PRODUCTION

I began with a preliminary reading of the book, which took place during a stretch of unemployment at the local coffee shop in a sort of public performance. Over 32 days I read 50 pages a day, drawing boxes around words to include and crossing out excluded words, such as homonyms, abbreviations and certain proper nouns. The mark making would ensure I didn't miss words during the actual recording, and this reading also served to teach me the phonetic pronunciation key.

I then sat at the computer and, for 32 days, 50 pages per day, recorded, edited and archived to cd the words. I recorded in 10 page batches, and read straight through. If I was aware of a mispronunciation, I paused, said the word 'no', and carried on. When the batch was complete, I scanned the SoundEdit 16 timeline to visually identify the errors and delete them. The word 'no' is the only word I am able to identify, to read, in waveform.

To create the books - each of which average 500 pages - it was necessary to make image files based on the waveforms. The only way I could figure out how to do this was through manual screen-grab: 72 days. I also needed a matching wordlist, but Oxford Canada was unable or unwilling to sell me a digital file. Again, I reverted to a manual process: 21 days. With the assistance of Ryan Johnston, the sysadmin at The Banff Centre, a perlscript was written to automate the naming of image files according to the wordlist, and an html template was used to automate the layout of the pages of the books.

The printing was done on an office Xerox laser printer, which ran 16 hours a day for 6 weeks.

The binding was done professionally by Colin Bate Books in Calgary.

The website firstly required the individuation of the batches of audio files into discrete files for each word. Using automated export and formatting functions in ProTools audio editing software in combination with manually creating regions for each word allowed me to create generically named files to which I applied another naming perlscript, written by Leif H. Askeland.

The artist and programmer Tom Leonhardt was contracted to write and setup the php for the mysql database, and setup Apache on a standalone imac. The process by which Tom and I defined the logic of the piece was incredible[III]. Over several weeks we generated many hundreds of emails full of question and answer, defining main goals, alternatives and do or die scenarios.

One of the most elegant developments came in finding a solution for the problem posed by users who searched for words not in my dictionary. We found something called Soundex [3], which is a

[II] http://www.usask.ca/art/horowitz_gbam/home.html

[III] I'm proud that this process also enabled me to learn a little bit about unix and navigating the command line!

form of coded abbreviations for names created for an American census about 100 years ago, and which is supported by mysql. Based on the coded abbreviations, **melitzah** provides nearest matches for users.

Then, Tom connected to my computer from Toronto - several thousand kilometers away - and did the work. He later told me that one of the reasons he decided to work with me on the project remotely was, simply put, because it was possible.

This systematic approach, the reading, recording, archiving, managing, saving, planning, typing, printing, binding, individuating, script-writing, server-side automating, collaborating, backtracking, formatting: it proved a particular form of digital materiality for me, and confirmed my belief that computer use does not always support the Cartesian duality so often applied to it. Those folders and files became objects that had as many perceived physical properties as a book in my hands or scrabble tiles on my deck. They did not, alternately, improve my scores or my ability to do crosswords.

The repetition of actions involved in making this work made me aware of the distinction between meaning and data or information. Information, in this instance, was more often processed without a direct associated meaning. I felt like I was inverting the proposition Roland Barthes made about the photograph as a message without a code: that I was working with a tremendous amount of code with no message.[2]

4. INTERACTIVE ART AND MELITZAH

Artists and curators often speak romantically about how interactive art enables and empowers users, gives the user complete control over the narrative or the proceedings of an interactive artwork. Interactive art is no more or less empowering than a choose-your-own-ending novel: there are variables, but they are limited. Interactive artworks are limited by physical and programming parameters, by media, by environment, and, in the case of **melitzah**, by vocabulary. The limits of interactive art are, in fact, a control of the user in the same way that smart graphic designers lead an eye across a page.

melitzah is limited by the possible combinations of the 65,835 words in the database, bandwidth, and the vocabulary of the user.

Despite this control, **melitzah** does enable in its own way. Viewers have the liberty to become creative. They are competent: they know how to conduct an internet search, and certainly know how to construct sentences. Users of **melitzah** [IV] can and do acquire a sense of authority and authorship over their interaction with the work, and they appear to take great pleasure in their discovery and play with the machine. They are satisfied when the words they seek are found, and alternately frustrated and tickled-pink when the Soundex code provides them with unexpected results.

[IV] In a way, these users are actors, they perform the piece.

5. CONCLUSION

melitzah is an art project that mixes up the visual relationship between language and speech by questioning the semantics of the symbols that we use. In aiming to make a closer connection between the words I speak to communicate my thoughts - in taking at face value the desire to make a science of semiotics - **melitzah** presents an oddly literal artificial system of semiosis. Its mimetic basis points, above all, to my belief that art is a communicative discipline and engagement, the purpose of which is, simply, to convey.

6. REFERENCES

[1] Barber, Katherine (ed.). The Canadian Oxford Dictionary. Oxford University Press. Toronto, 1998. 1602.

[2] Barthes, Roland. Image Music Text. trans. Stephen Heath. "The Photographic Message." Noonday Press, Farrar, Straus and Girous. New York, 1993. 17.

[3] Moore, Karl. The Secret of Soundex. http://www.developer.com/net/vb/article.php/10926_1539851_2. 2002. [entire article].

[4] Scherman, Rabbi Nosson (ed.). The Chumash: The Torah, haftaros and Five Megillas with a Commentary Anthologised from the Rabbinic Writings. Mesorah Heritage Museum. Brooklyn, NY, 2000

From Text To Interface
Theatre And Digital Media

Wolf Ka

res publica
13, rue Germain Plion
75018 Paris, France
+33 1 42 62 29 03

compagnie_respublica@yahoo.fr

ABSTRACT

The following text by the digital performing arts company res publica, confronts theatre from a practical perspective with an approach informed by theory of communication.

The first part of the article introduces res publica's point of view concerning performing art. This serves as a methodological framework for the analysis of the recent work "Enjeux" (Inplay), an interactive scenic device.

Keywords

Scenic device, semiotics, digital scenography, scenic interface, hypertext, interaction

General Terms

Documentation, dance, theatre, digital media, technology

1. INTRODUCTION

The company res publica is concerned with the stage as space of representation. Their performances precede neither from a body-oriented standpoint (dance) nor a dramatic narrative oriented one (theatre). Their investigations of stage question the construction and arrangement of space, the structure of representation and the relationship of stage and public. Their centres of interest are formalized in "scenic devices".

The term of the "scenic device" is defined in opposition to concepts of representation in terms of a metaphorical translation of reality. The creations of res publica are scenic devices where the theatre as place of performance as such, is analyzed by the transformation of its setting. The performance is understood here in the sense of a cultural practice and thus socially, historically and technologically determined. In this perspective, the devices both deconstruct and reconstruct representation. The scenic devices operate by the categories of space, interpretation, public, time, and action.

In an historical perspective the work of Oskar Schlemmer, the founder of the stage at the Bauhaus in Dessau, is an important reference for the work of res publica and in particular for the creation "Enjeux".

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

He was one of the first theatre artists who worked in a systematic and radical way on the human body as a code, and the relation of Man and Space. He proposed a perspective of body on stage as an abstract and symbolic representation as opposed to the psychological and expressive approaches of German Dance in the 1920th. He intended in his stage work to represent the exemplary, the universal instead of showing the individual or the typical. His theatre works did not want to be a copy of reality but an artificial (art) product in order to bring up the essential, and the (pure) idea. The term of "mobile Raumplastik" (mobile sculpture) illustrates his concept of human body representation.

He examined the gesture/ movement as a relation of Man and Space on an abstract level and not as a representation of human psychology. The result is the formalization of human motion to geometrical forms. He investigated the stage in his masterwork " Triadische Ballet" to study the relation of space and body; through the basic (triad) parameters of form-colour-space, height-depth-width, and the three basic forms of geometry: sphere-cone-cube.

He opposes the laws of the cubic space of stage to the laws of natural man. If the space is adapted to man, the stage becomes naturalistic or illusionist. If man is adapted to the cubic space the stage becomes abstract. From his viewpoint, the laws of the abstract stage are the invisible lines of planimetric and sterometric relationships. [3]

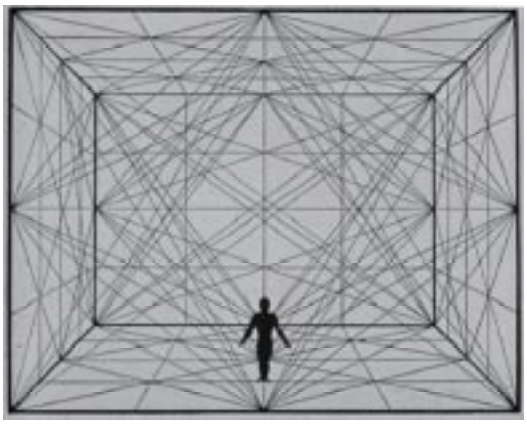


Figure 1. Oskar Schlemmer, "Figur und Raumlineatur" (Figure and spatial delineations), 1924

Even if Schlemmer didn't included technologies in his stage works, his work is an important source in the history of digital performance since he was studying the relation of different media in the context of stage, without any metaphorical approaches. [2]

The idea of the scenic device, as well as the question of the integrating digital media and the consequential re-organization of the semiotic situation in the performance process will be discussed here with examples from the performance "Enjeux". Furthermore, it would like to point out the value of both dance and theatre in the context of performance where digital media are integrated.

2. PAPER

The French company "res publica" developed the interactive scenic device "Enjeu3+4x3" between 2001-2003 in collaboration with the Belgian office for architecture and urbanismLab[au]. At the centre of the scenic device is an interface, a multimedia front projection onto a screen at the back of the stage. The projection is at the same time the only source of light and the only scenography. In this way, the projection is not just a visual representation but also a complex device that plays on the space in the projection's light cone. The interface relates the screen projection to the space of stage. The public manipulates a trackball interface and consequently they also alter the space of representation.

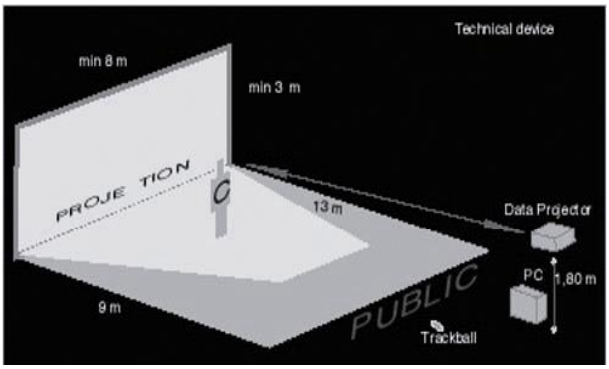


Figure 2. Scheme of the technical device.

From the beginning of the performance one trackball circulates in the audience. Since the trackball is lit up and the movement of the cursor is always visible on the screen, the public can easily interact and participate at the action at any moment. In the beginning of the performance, the public just choose a sequence by a menu and progressively manipulate the whole sequence.

The creation shifts the status of the public from a group of individuals to an assembly, integrated in the construction of the content and the construction of the representation.

In order to evaluate this idea and the idea of a hyper-textual dramaturgy of representation, a meta-text layer is integrated into the architecture of the interface. After each sequence the meta-text layer appears. Every title of the played sequences is written in horizontal from left to right on the screen. In this way, the public sees the choice they have made.

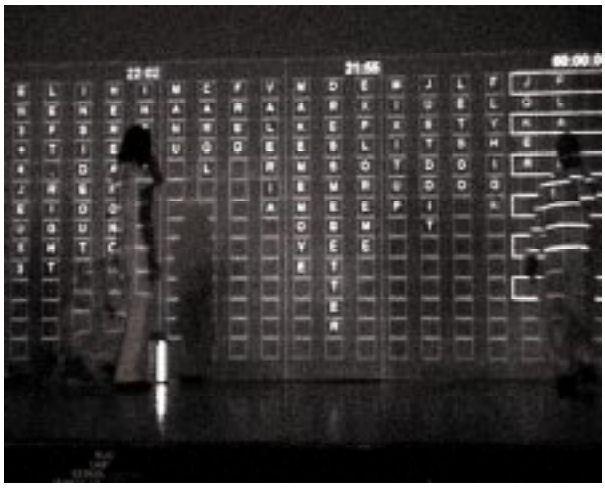


Figure 3. Meta-text layer at the end of the representation

From a theatre-oriented perspective the interface has the status of a text. The shift of a linear to a hypertext structure, the extension of a mono-semiotic vocabulary (alphabet) to a poly-semiotic vocabulary (alphabet, graphics, picture, sound etc.) and the autonomous presence of the text in the process of representation, that is independent from mediation of the dancer (as opposed to a dramatic text or a partition of dance or music) are the guiding principles in the theoretical approach of "Enjeux".

This scenic device questions the relation of two different systems of signs: the graphical interface and the performers representation. The iconographical language of the interface and its consideration as an architecture is necessary to establish an interaction between stage and public. The relevance of the overlay of the two systems (graphic and performance) depends on the ability of the signs to institute codifications. The interface refers to graphic, linguistic, iconographic signs and the second system refers to the expressive abilities of the human body. The purpose is to link these two systems in order to allow the public to intervene in the process of representation. This means that the question of 'textuality' in this performance is based on the creation of codes that are proper to this multimedia-light space and to assign them to codes of human expression.[1]

The dramaturgy is based of the stage-specific mediated situation, as a spatially and temporal limited action carried by performers. The categories of space (where), performer (who) and action (what) compose a three-part dramaturgy. Each part is composed by several sequences, which then develop into their own optional or hyper-textual dramaturgy.

In the following a short introduction to each of the three parts institute the general idea that are illustrated by two sequences.

2.1 Part A - WHERE

The sequences of part A, introduces elementary codes based on the light cone, the screen, the stage and the performer. A vocabulary is introduced, referring to lexical, graphical and iconographic codes, which create spatial situations in order to evoke imaginary narratives and physical settings. Thus, the inter-play of these various code-languages within a sequence is used to

construct a situation, an interactive scenic space and introduces of our first narrative: WHERE.

The public has the option to choose four of six sequences by a menu on the interface. Each title of sequence refers to an idea of space questioning the complex relation between the representational and concrete situation of the stage. After the fourth sequence the interface skip to part B. The fourth sequence that is selected will define the vocabulary for Part C.

2.1.1 Part A example "In transit"

The projection of the word 'airport' introduces a place and a situation influencing the audience's reading of the stage. The sequence is based on a play of associations with the word 'airport'. When the dancer's bodies intercept the projected letters, the reading of the word changes, as does the place and situation it suggests. This inference continuously constructs and deconstructs the scenic space between the material space of projection and a suggested one.

Using a "spatial language" (choreography) the dancer writes new semantic connections to the "space of departure" by using his body as a medium. In this way the public reads one letter after the other of P-A-R-T-I-R (to leave in French), composed with the letters of A-I-R-P-O-R-T. The letters on the bodies travel from the ramp to the screen and translate the semantic content of the word into a choreographic language. Later, the dancers "write" the words "Rio" and "Porto". The sequence plays with the imaginary and material situation of the stage and sets it in relation to a linguistic and chorographical code.

Figure 4. Sequence In transit : "Rio" on the body of the dancers

2.1.2 Part A example "in front of"

The sequence "In front of" sets the space of the stage in relation to the performer using a graphic code. The spatial construct of the sequence is based on a projection of animation composed by RGB coloured rectangles and their synthesis (white). The interface projects rectangles at a regular frequency of 10 seconds in four different sizes and in the four colours (RGB and white). The bodies of the Performer intercept the rectangles. Each size of rectangle corresponds to a position in the space, and a sound of a 4-tone chord. The smallest rectangle is reflected directly at the screen, the largest at the ramp and the two furthest at a distance in between, in order to reflect the same size of the rectangle. Because of the diagonal cone of projection the bodies intercept the light at different heights, depending on the distance to the projector (falling onto the head, shoulder, belly and hipbone). Each colour corresponds to a body position such as front on, back, profile left and right. From this codification that incorporates space, colour, form and parts of body, the stage becomes a choreographic play of 4x4x4x4 units, comparable to a Mondrian human-space composition.

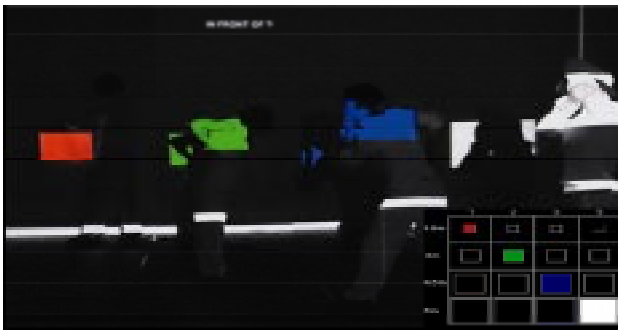


Figure 5. Sequence "In front of" with scheme of projection

2.2 Part B - WHO

The second part of the performance examines the status of the performer on scene – "WHO", in relation to their representation on stage and on the screen/interface. While in part A the four performers have been identified through denoted codes: numbers and positions (for example, Number 1 - Position 1 – Colour red), the second part constructs a specific association to the performers (a connoted code) through distinct actions and interactions, thus shifting the notion of figurant to the one of performer.

The relation between the performer, the stage and the action is realized according to visual communication codes (both connoted and denoted) and the visual language moves between icons, symbols and images - figurative or non-figurative codes.

The second part establishes specific relations between the public and the four performers. A direct interrelation is used to enlarge the interaction between the public and the stage, from choosing different sequences to interacting inside the sequences with the performers.

The overall concept of the second part: "WHO" allows the public to choose between four non-interactive solos, introducing a specific code of each performer. After the first solo, the public has the option to explore these specific codes by playing interactively with them. By choosing "No" the second solo of the performer is removed.

2.2.1 Part B example „Number 3" (simple solo)

The sequence confronts a scenic representation with a textual one through the projection of the sentence 'JE JOUE MANU' (I play Manu) in white letters on black background. While moving backward in the projection's light cone, the actor intercepts the letters of the sentence on his body, inscribing on it "MANU" / "MAN" / "NU" (naked in French) whilst successively taking off an item of clothing. In this manner he transforms the sense of the

sentence from 'Je joue Manu' (I play Manu) to 'Je joue Man' (I play man) and - 'Je joue nu' (I play naked) and yet the performer stays, through his actions, coherent with the sentence written on the stage.

Figure 6. Sequence "Number 3"

Does the actor play Manu - does he play a man - does he play naked or is he a man whose name is Manu who is naked on the scene? By the use of this textual combination the status of play in scenic space as one of action and representation is formalized - a relation which according to the introduction of interactivity gets at once more open because of the direct relation established between the public and the action/representation of the actor.

2.2.2 Part B, example "Make me moving"(interactive solo)

The public interacts with the dancer through a schematic but articulated graphic of a skeleton, projected on the screen, where it's manipulation leads to a successive shift from a figurative to an abstract representation.

The dancer directly transposes the public's interaction into body movements but in addition the dancer transposes more abstract ideas such as flexibility, lightness, speed. The dancers movements always depend upon the graphic's distortion and the way the public is manipulating the skeleton. The dancer worked out during the rehearsals different approaches of translating the graphic distortion into a chorographical language. As there are infinite possibilities the way the public act, the action/representation is improvised.

The cartographic mapping of the manipulations / movements thus becomes the result of establishing a relation between the public and the dancer, introducing 'dance' not only as a body moving in space but also its relation to gesture (figurative - physical) and composition patterns (geometry - abstract). In this way, the sequence enlarges a textual and graphical vocabulary of interactivity to include gesture, expression and movement.

Figure 7. Sequence "Make me moving"

2.3 Part C - WHAT

The PART C examines the action on scene - "WHAT" - through the space of interaction and the interpretation. While intervening directly on the projection/ scenography, the audience determines the sequence, the duration of the sequences and the modalities of interpretation. Each sequence explores the space and interpretation, parameters introduced in PART A and the PART B. The interpretation of the sequences of PART C is coded, but their realization remains to the result of the public's intervention. The end of the PART C is conditioned by the duration of the performance. Once this time is passed, the public can decide to continue the performance, in which case, the sequences of part C are again proposed, until the spectator chooses the final "no".

2.3.1 Part C example "Mix it up"

The sequence "Mix it up" is based on the same code as the sequence "In front of". Four rectangles in the elementary colours and white establish a basic vocabulary of postures, the dancers arms forming a geometrical correlation to the projected rectangle. (Red = diagonals from left to right, green = horizontal, blue = diagonal = from right to left and white = vertical.) The size of the rectangles determines the distance to the screen.

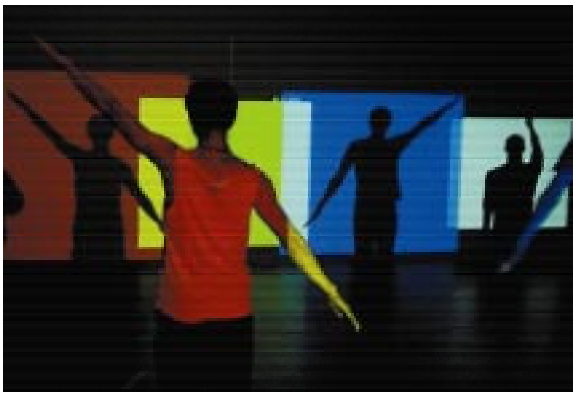


Figure 8. Sequence "Mixitup"

In contrast to the sequence from part A where the changes of the rectangle was a written-partition, the animation in part C depends on the position of the cursor on the axes x and y. Each

rectangle behaves differently concerning the changed position of the cursor. When the rectangles are overlapped, fields of its colour-synthesis appear and give the dancer the possibility to change the colour field. In this way the movements from left to right of stage depend more or less of the choice of the dancer, the movement from front to back depend on the publics manipulation. The codification of colour, space and body permits a real-time notation of choreography formalized by a dynamic light-scenography. The choreography is the common result of public and dancer over the codification of space and body by colour and form.

2.3.2 Part C example "Playball"

In the corresponding sequence of part A “left-right“ a small white dot crosses the black screen from left to right giving the illusion of a tennis ball where the graphic layout refers to the one of the first video games “Pong”. The same graphic sign is used in the sequence "Playball", whereby the ball movement is this time not programmed but is caused by the cursor movement manipulated by the public. Left or right movement of the mouse moves the light cone on the horizontal axe, the movement upward or down increased and/or makes the light cone smaller. The public is invited to play and this possibility is introduced by the title: Playball = play with the ball/ performer. In the sequence of A the light cone produced the illusion of a ball. In the sequence "Playball" the relation of material and illusionary space is reversed. The (play) ball becomes on its semantic level a light spot, which is what it is.

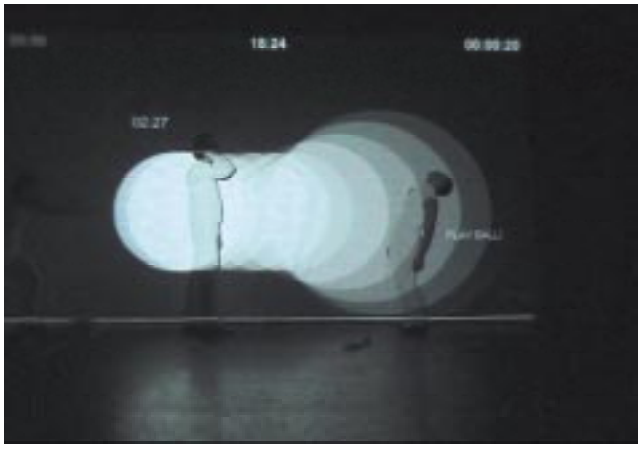


Figure 9. Sequence “Playball”

The play of the sequence consists of the performers moving in the dark stage and taking positions related to past sequences. As soon as the light cone lights them up, the performer freezes to a picture. The resulted "flash backs" re-contextualize past situations, positions, gestures, etc to new semantic associations which are proposed by the performers and selected and cut by the public.

3. CONCLUSION

The performance “Enjeux” and this theoretical reflection, aimed to contribute to the understanding of how technologies modify cultural practice by integrating an interface in a communicational situation within theatre representation. The analysis of “Enjeux” deconstructed a textual based representation and illustrated how the integration of digital media modifies the status of text, performer and public. It illustrated that the integration of new (or simply other) media provokes new plays or modalities of representation.

Furthermore, this perspective of theatre attempts to demonstrate that a semiotic approach to theatre in both its theoretical and practical dimension could be fruitful without literature based or narrative approaches. The text wants to propose a critical view in the field of theatre and media, by focusing on the space of representation as a category for theatre analysis. The text questions scenic representation not as a discipline of theatre or dance, and integrates body and text (and media) into a non-hierarchical relation.

The analyse of the scenic devices within “Enjeux” aims also to document that the “digital paradigm” could be a collective experience in a physical space instead of an individual experience in front of a computer screen. By integrating the public in the constitution of the representation, the work questions also the social dimension of digital media within the symbolic level of theatre. The shift from a public assisting a performance to a public that participates by manipulating the basic parameters of representation becomes especially crucial at the moment when members of the public have to decide to end of the representation. As it is only one person who takes the responsibility for the rest of the public to continue or to end the performance, there is often a moment of negotiation where the spectators starts to express their agreement or disagreement. By moving the cursor from yes to no and back to yes the person with the trackball plays with these reactions . In this moment the communicational situation alters to a play between the one with the trackball and the others, but without ever leaving the framework of negotiating the condition of theatre representation.

Figure 10. Interface “The end”

4. ACKNOWLEDGMENTS

The theoretical and practical approaches of the creation “Enjeux” was developed together with the laboratory for architecture end urbanism Lab[au]. Parts of this text was written by Manuel Abendroth of Lab[au] and published on the website of Enjeux www.lab-au.com/enjeux.

I wish to thank Helen Evans for the review and correction of this article.

5. REFERENCES

[1] About Textuality, theatre and new media: Bauchard, F. Le théâtre et les nouvelles technologies. In éc/arts, publishing site, 2002, 160-165.

[2] About Schlemmer as avant-garde of the notion of digital space: Norman, S. J. Corps / espaces interactifs. in___ 156-165.

[3] Schlemmer in general: Schlemmer, Raman C, Oskar Schlemmer, Musées de Marseille reunion des Musée nationaux Musée Cantini, 1999.

A Recombinant Information Space

Andruid Kerne, Vikram Sundaram
Interface Ecology Lab | Center for Digital Libraries
Texas A&M Department of Computer Science
College Station. TX, 77843. USA.
+1 979.862.3217

{andruid,vikrams}@cs.tamu.edu

ABSTRACT

The traditional web browsing paradigm is one of the most significant developments in the histories of publishing and human computer interaction. Yet clicking through links grows tedious, without affording dynamic knowledge building. What's needed is a partially ambient visualization mechanism that sifts through focused subsets of web space, with our occasional attention. CollageMachine supports fluid adaptive dynamic browsing. The granularity of browsing is shifted down from documents to their constituent media elements. As time passes, elements of interest stream continuously into a recombinant information space that helps us locate and arrange information, and generate connections. An interactive interface maps the expression of interest together with direct manipulation of visual design. This mapping enables the user to effect the same information space visualization that the program is developing, and express dis/interest, with a single gesture.

CollageMachine's generative approach to browsing extends experience beyond routine, to support creative experience. Stochastic decision-making enables the recombinant information space to develop in a way that is predictable on the average, and yet still open to the unexpected, as while browsing a physical library's stacks or store's racks. Hypermedia structure forms the basis for an associative model that interprets the participant's actions; like human memory, it learns through spreading activation.

A diverse set of techniques, that have evolved through several generations of development, make this work. *Seeds* are initial documents that are fed to the collaging engine. A document's HTML markup serves as a contextual guide for breaking it down into information elements. Image processing creates cohesion among the elements, and foreground/background relationships that make the collage easier to read.

Keywords

recombinant media, recombinant information, web reassemblage, granularity of browsing, compositing, metadocuments, navigation, recommender systems interfaces, creative cognition

1. INTRODUCTION

The traditional web browsing paradigm is one of the most significant developments in the histories of publishing and human computer interaction. With hundreds of millions of users accessing billions of hyperlinked documents [36], something is clearly working. Yet clicking through links can grow tedious, especially when you're not exactly sure what you're looking for, or where to find it. What's needed is a mechanism that lets us sift through focused subsets of web space, in a way that is stimulating and enjoyable, but not demanding. Such a mechanism can be directed when we want to interact with it, and can still work for us while we're busy with other activities. It creates visualizations that we focus our attention on sometimes, and leave in the peripherae at others. Links are traversed automatically, by a web crawling interface agent that models our interests, and acts accordingly. Using its model of user interests,

Figure 1. Two states from a recombinant knowledge space session, featuring research and student work in visualization and computer science at Texas A&M.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group
Copyright 2003 by Andruid Kerne.

CollageMachine automatically retrieves relevant information, and visualizes it in a mutable information space. This retrieval and visualization is a form of automatic browsing.

2. RECOMBINANT INFORMATION

Recombination is the process of taking existing coded compositions, breaking them down into constituent elements, and recombining those elements to form new codings. When practiced with and by living organisms, the constituent elements are genes, which contain essential information that defines an organism's makeup. Shuffling of base pairs, mutation, and splicing of subsequences are recombinatorial means that produce new gene configurations, and from them, new beings. Genes are the code of life, from which new organisms are created, and through which new species evolve. The information content of genes is so significant that the field of bioinformatics has emerged to focus the application of computational methods to gene composition and recombination.

Dada Information Collecting

In the information age, analogous processes of recombination are conducted with the units of meaning that we use to communicate, to express ourselves, and to represent ideas. The process began in the art world, in the Dada movement of the early 20th century. Marcel Duchamp, using the pseudonym R. Mutt, did the first creative work with found objects, or *readymades* [25]. He submitted "Fountain," an unadorned urinal, to the exhibition of the Society of Independent Artists Exhibition in 1917. The work created an uproar. People were outraged. Through this, he demonstrated the power of context in the interpretation of an object. Attention was brought to the object's semiotic makeup. Meaning was transformed by the shift from a bathroom to an art exhibit. The readymade *read* differently. The object's information content plays an essential role in its function.

According to an anonymous work attributed to Duchamp, "Mr. Mutt did not make the work. He CHOSE it [25]." This choosing is the act of collecting, reusing, and recontextualizing. It is the same choosing that contemporary D.J.s do when they assemble mixes. It is the same choosing that people do when they use file sharing services to swap musical tracks, and digital video recorders such as Tivo to collect video programming. Duchamp's work is not significant simply because of its effect on the art world. Duchamp asserted and demonstrated that "choosing," that is, collecting, is a creative act.

Duchamp's Dada associate, Max Ernst, used found objects as the genes of new visual semiotic forms. In his collages and overpaintings, found elements are combined with new ones. Collage literally means, 'put together with glue." Ernst's work followed that of *papier colle* artists such as Picasso, who, in one work, substituted a physical spool of thread for a painted one. Yet their contemporaries, Tristan Tzara and Louis Aragon [3], asserted that Ernst invented collage several years later. This was, again, because of semiotics, because of the emphasis on meaning and concept in the construction of Ernst's works, and the way that found objects are used to produce new meanings. For example, as Krauss so clearly analyzes [23], *The Master's Bedroom* is an overpainting of a page from a catalog of high school teaching aids. Ernst overpainted most of the sheet with a simple plank floor and walls. He leaves an oversized bear in the distant back of the room, where the walls are converging, and a tiny whale in the front. A paradox is created, in which the rules of perspective are contradicted. Thus, he uses the catalogue in a commentary that challenges modernism, through its alignment with perspective. The precise choosing and the overpainting combine to a create an evocative new species of meaning.

Practices of Authoring by Reference

As the information age has developed, similar recombinant practices have been conducted in spheres of art, entertainment, and information systems. Duchamp's friend, John Cage, used found objects in music composition. For example, in Imaginary Landscapes #4, the settings of the dials of radios are scored for live players (1954) [33]. A few years later, Stockhausen's *music concrete* was the first work in which tape recordings of environmental sounds played an essential role in compositions [39]. In the conceptual art movement that began in the sixties, ideas themselves are considered to be an artistic medium [26]. Now, in popular music, hip hop artists quote sounds iconographically. For example, in the early nineties, Public Enemy utilized clearly quoted samples from television and Malcolm X, as well as James Brown [32]. They are among legions of practitioners. Knowledge of art history is not a prerequisite for engagement in recombinant practices of cut and paste transformational remixing.

With the proliferation of the networked personal computer, collecting becomes an everyday process that permeates work, communication, and browsing. Vanevar Bush foresaw the potential for this in information systems in 1945 [6], thirty years after Duchamp and Ernst began the practice with mass-produced objects. The trails of the Memex are based on an associative process of collecting and annotating. These annotations, which represent the collecting individual's responses to found elements, have the same creative potential as Ernst's overpainting activities. Bush foresaw that trails, themselves, would function as a medium, and that authoring of them would be an important activity. Soon after the onset of digital computation and communication, Ted Nelson formulated the concept of hypertext, a mechanism for non-linear writing, based on the reuse of found textual objects, and the power of random access memory to support the structure of the reference [29]. With the advent of the Internet, and WWW technology, a form of hypertext authoring and browsing became the fastest growing publishing medium in history. Thus, we experience a convergence of practices of *authoring by reference*. Explicitly named information systems intersect with information-based art and entertainment.

The work of Duchamp and Ernst is a beginning of the information age, because it focused attention on the meanings of objects. It brought focus onto how they *read*. This work started the postmodern era, because it brought production in the art world away from the creation of new masterpieces. Instead, work is created through the recombination of existing objects, based on their semiotic and sensory functions, through choosing, annotating, and assembling.

Media are sensory forms that information takes. By recombinant media, then, we mean media that is created through the combination and composition of preexisting readymade media elements. One definition of information is *data that communicates*, and a definition of communication is the *transmission of meaning*. We use media to convey meaning. Thus, notions of information and media are inherently connected. The distinctions lie only in emphasis on sensory

Figure 2. Snapshot from a 9-11 recombinant information space browsing session.

forms or communicated ideas, which cannot be essentially separated. By recombinant information, we mean meanings created through composition with readymade information elements.

Deep in the information age, we are deluged with information, and wading through it. There is a shift in the experiences of a broader segment of the population. We have a growing need to collect, arrange, organize, and assemble readymades. We need to make sense of these collections, to keep track of things, to understand interrelationships. The forms include bookmarks, email attachments, and more explicit *metadocuments* [11]. The contexts of signifying acts of choosing move from the public art exhibition to the personal computer and the Internet..

Recombinant Metadocuments

Metadocuments are authored by reference. That is, they are documents that consist primarily of references to other documents, including elements from those documents, and annotations. Users of hypertext often need to collect references to significant places that they encounter while browsing. These collections are metadocuments. They consist both of references by name, such as `<img` tags and hyperlinks, and by value, in the form of textual quotations. We call the image references and quotations *information elements*. For each information element, in addition to any embedded hyperlinks, there is always an implicit reference back to the original document, which we call its *container*. Schraefel articulates the importance to metadocument authors of the connection between an information element and its container [35]; when we collect information elements from the web, we want to be able to easily return to the sources of the quotations. Each information element can be thought of as a fat bookmark. When they include information elements, and the ability to navigate back to containers, as well as over to hyperlinks, metadocuments make referential structure functionally explicit.

Prior systems for metadocument authoring, such as Walden's Paths [11] and VKB [37], have focused attention on work practice. With its support for *spatial hypertext*, VKB even addresses the work practices in which people use color and font characteristics to group information elements. Still, these programs have not focused on the creative recombinant potential of collections of visual semiotic elements. Our senses of creative and work-based properties and uses of collections of information elements may come from different spheres of human activity, and different processes of production. Yet, they are essentially compatible. People tend to collect stuff that is interesting to them. We expect that for the collections, themselves, to be interesting to look at, will be beneficial. Metadocuments can transform their constituent elements through the recombination of information. Users do not have to compose with the same degree of precision as Ernst, in order to create compositions *they* find meaningful. Recombinant metadocuments are rich potential

sources of evocative, expressive new media. CollageMachine is a process-oriented art work, and a creatively motivated tool, in which the activities of browsing the web and authoring metadocuments are integrated.

Constructive Operations of Recombination

Assemblages connect found elements. They build relationships between the elements, and invite processes of interpretation. In this way, they form information spaces. As recombination moves from physical to digital forms, the potential grows for recombinant information spaces to be built procedurally, and to evolve dynamically. Whether compositions are single state and static or procedural and dynamic, whether they are composed with physical objects, or digital ones, certain underlying mechanics remain consistent. Whether it is performed by a human, by a computer program, or by a combination thereof, the process of creating recombinant media consists of four fundamental constructive operations:

Selection – Choosing material is typically a two pass process. First, one or more sets of candidate media elements are collected. In Max Ernst's work, this involved identifying scientific and popular catalogues (documents), and then selecting images from them (cutting). In CollageMachine, it means downloading and parsing documents, and building one set of candidate document references and another of media elements. Then a second pass of selection decides, periodically, which candidate web document references to crawl via recursive re-invocation of pass 1, and which candidate media elements to include in the current visualization.

Spatial Arrangement – deciding where to place elements spatially, in relationship to each other. In the digital realm, we can scale images and text, so this also includes determining elements' sizes.

Treatments – processing of individual elements in the recombinant composition. Rauschenberg, for example, sometimes puts a layer of varnish or glue over an image, so as to dull or brighten its appearance [15]. In *His Master's Bedroom*, Ernst's overpainting only partially obscures the material underneath it, leaving a murky echo of the obscured elements, that suggests the process of memories drifting into the unconscious [23].

Digital treatments include filters like blur, down-sampling, hue, saturation, and value re-mapping, and Fourier resynthesis. In CollageMachine, images are normally placed without initial treatments. Over time, as the age on screen, they are desaturated. They may be blurred as a reflection of the user's expression of negative interaction. The exception to this is when the model of the participant's interests maps negative interest to an element before it is placed. In such cases, the elements is blurred to begin with.

Fastening – the means of assembly; processing that pastes, blends, morphs, composites, or otherwise connects elements. We must answer such questions as: are elements simply juxtaposed, or are they blended further? Are lines of attachment softened, or do they stick out? In *Reves et Hallucinations*, Max Ernst leaves visible pasting lines [1]. In *His Master's Bedroom*, he makes them invisible [23].

One digital means of fastening, which is invoked particularly in the world of video, is the alpha channel. Alpha affects transparency, and, through intermediate values, translucence. When layers of bits overlap, the alpha channel becomes a fastening technology. A technique, known sometimes as alpha masking, uses intermediate alpha values in the border region of an image. Implemented in CollageMachine in 2002, such alpha masks are used to create a sense of visual flow between elements. They are implemented through the procedural generation of an alpha gradient that runs from an extreme alpha value at or near transparent at the image's edges, to absolute opacity, as the mask area ends toward the center of the image. The resulting sense of visual integration is palpable. The conceptual tensions between elements are often strengthened by their visual integration. Other elements are fastened opaquely, retaining separation. Prior versions of the program, which relied exclusively on juxtaposition, were less effective in conveying semiotic sense. The addition of image processing algorithms, of which alpha gradients are one, and buffering techniques which allow these computationally expensive operations to be performed gradually, without slowing the program down too much, did not require rewriting the entire program. They are modules in the CollageMachine recombinant framework.

Manovich focuses on the properties of fastening technologies [27]. Because montage and collage leave hard lines between composed elements, where alpha masking and morphing create continuity, he calls compositing "anti-montage." At the same time, he identifies montage as a device for creating conceptual tension, through juxtaposition between disparate elements. This makes sense semiotically. Yet this is not inherent in the media of fastening. It is a property of information composition. Manovich's examples of how actors are shot in front of blue screens, and then seamlessly composited into archival or computer generated backgrounds, clearly describe typical uses of digital compositing capabilities. Yet, as current versions of CollageMachine demonstrate, the same technology can be used to create conceptual juxtaposition. Collage, montage, alpha blending, remixing, and morphing are genuses of recombinant fastening operations. The compositional dimension of selecting such techniques is relatively independent of the dimensions in which visual and semiotic choices are made to create more or less cohesion instead of heterogeneity, in a composition, In a generative recombinant system, there is a whole space of options for how to integrate modules which perform semantic and visual analysis, with those that generate visual composition.

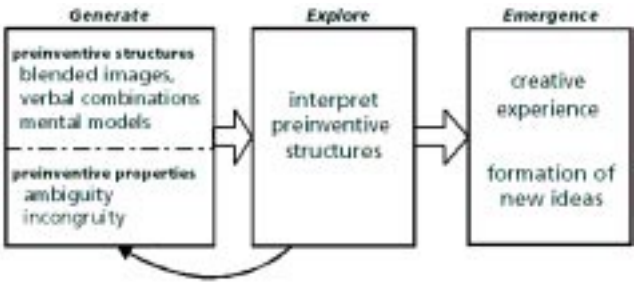


Figure 3. Creative cognition: the Geneplore model.

Cognition of Recombinant Information

Our cognitive processing of recombinant information is addressed by the *geneplore* model of creative cognition [10]. According to geneplore, creative experiences sometimes develop when phases of generative processes (e.g., memory retrieval, analogical transfer) alternate with exploratory

interpretive operations (e.g., attribute finding, hypothesis testing). Certain conditions increase the likelihood of creative experience. The generation of *preinventive structures*, which serve as the grist of creative process, makes the development of creative results more likely. Combinations of images and words, that is recombinant information, are a form of preinventive structure, as are visual patterns and mental models. The exploration phase consists of articulation, interpretation, and refinement. We play with the preinventive structures in search of understanding. We may iteratively cycle back and forth between phases of generate and explore.

Some preinventive structures are also characterized by *preinventive properties*. Examples of these include ambiguity and incongruity. That is, when information elements are recombined, if the combinations make sense immediately, the cognitive process is not likely to go anywhere. But, if there are potential relationships that are not immediately clear, the mind tends to work on making sense of them, to find new connections. Sometimes, configurations of preinventive structures don't lead anywhere. There are no guarantees. On the other occasions, we experience, "Ah-ha!" This is the emergence of new ideas.

In the case of a recombinant work with fixed form, such as a still collage, or a filmic montage, the creative process actually includes two phases of exploration. First, there is process of the artist, making the work. The artist is privileged by the ability to generate forms of the work, reflect on states as the work develops, and iterate back through more creative generation. Eventually, the work is presented to an audience. Here, again, there is a new process of interpretation. Typically, the conceptual relationships between elements of a provocative work, such as Ernst's collages, are not immediately apparent. They are preinventive structures with preinventive properties. The audience member has an opportunity to engage in interpretation that may lead to emergent ideas. The audience member is not, in such cases, able to influence the generative phase.

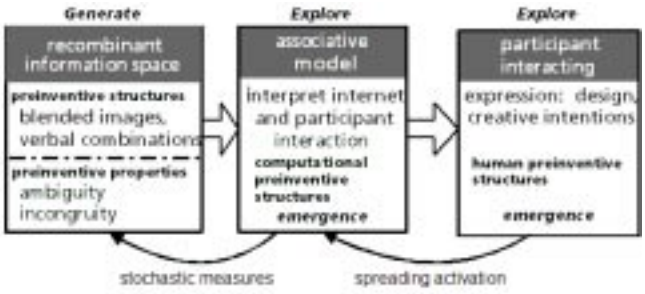


Figure 4. Creative cognition with CollageMachine.

CollageMachine is being designed to create a more actively *geneplorative* experience. The difference is that generation of recombinant forms, as well as their interpretation, continues once the participant user is involved.. The computer plays the generative role, by undertaking procedural dynamic retrieval of information elements, selection of them, and recombinant visual composition. Exploration is accomplished cooperatively both by the participant, and by the program. The participant can use the interface interactively to express interests. The program's goal is to translate this interaction into responsive automatic browsing, that is consistent with what the participant is actually expressing. That is, the goal is to retrieve related elements to those that s/he has expressed interest in. The program also seeks to recognize and respect the participant's design decisions as it makes spatial decisions about the placement of new elements. The program's mechanism for generative automatic browsing actions is its associative model of hypermedia content and user interests. Through this feedback mechanism, interactive exploration iteratively effects the generative phase.

Personalized Web Recombination

The web continually grows as a vast hypermedia repository of size beyond human scale. A small, yet precious, subset of this media is of interest to an individual. Helping people discover and derive value from that interesting subset is the essential role played by an emerging field of programs that perform dynamic recombination of web content as a reflection of a participant's personal desires. In Bender's "Daily Me," [5] – exemplified by systems such as FishWrap [8] and Zwrap [13] – the presentation takes the form of a "personalized newspaper." Anderson and Horvitz's "My Montage" [2], composes a similar presentation based on "routine patterns of access." These programs are influenced by recommender systems, such as Letiza [24] and Fab [4]. The impetus for another track of web recombinators generators, such as The Impermanence Agent [41], and Netomat [42], has come from the art world. CollageMachine, an agent of streaming web recombination that integrates functional and artistic motivations, has been under development since 1996 [17-22]. A burst of development during the last year has improved the agent model, added visual compositing, image processing, a language for specifying seeding, new methods for text visualization, a new interface design, and many other features. Integrated approaches to content retrieval, interactivity, and visualization are iteratively developed. The current conceptual approach and functionality set is a product of many design/implementation iterations over that period of time, based on diverse feedback through usability studies [16, 18], informal demos, public presentations, and conceptual walkthroughs [22]. The program has been co-developed through and along with the concept-context-design interactivity development model [22].

3. GRANULARITY OF BROWSING

Granularity refers to the size of the fundamental units that browsers present. CollageMachine shifts the granularity of browsing down from documents to smaller units. It makes *information elements* essential. In the current implementation, supported information elements consist of images and chunks of text.[1] A study by Schraefel et al [35] confirms that users regularly need to deal with these finer grains. CollageMachine gives attention both to the ongoing retrieval of interesting information elements, and to their visual composition into a recombinant form that facilitates reading individual elements and relationships between elements. This process develops gradually and continuously: pages are automatically streamed into the browser every few seconds; media elements once per second.[2] Thus, temporally, the granularity of browsing is

1 Support for additional MIME types, such as video, audio, and Flash, is in progress.

[2] These are the default; the user can halt content streaming, and control its rate (See below). It is also interrupted automatically by interactive media element drag operations.

extended from a single static view that settles and waits for the user to click a hyperlink, to a dynamic one that evolves continuously.

4. RECOMBINANT BROWSING AGENT

Conceptual And Functional Approaches

CollageMachine integrates three essential components: an associative model that drives operations, a recombinant visualization that reflects the state of the model to the user, and an expressive interactive interface that enables the user to directly manipulate the visualization and effect the model. As the program runs continuously, the model evolves. The model connects the structure of the information elements, including their interrelationships, with the user's interests and intentions. The state of the model at any moment is stochastically applied in all recombinant information-making decisions, such as the selection of documents to retrieve, of media elements to display, their size, and location. This effects dynamic adaptive browsing. An operation-specific metric is applied to the set of attributes that adorns each content element structure: each candidate is assigned a floating point weight.

Decisions based on the model are made probabilistically, rather than deterministically. Rather than choose a minimum or maximum, a weighted random select operation is conducted, These procedures, on the one hand, faithfully reflect the model, on the average. At the same time, by retrieving and displaying content elements that are within the scope of the traversed space, but are not necessarily the ideal candidate, CollageMachine opens the process of browsing. Scientifically, this works in a manner that is analogous to simulated annealing [9]. Experientially, it is like going to the library with a list of books and call numbers, and finding an exciting book on a nearby shelf; or finding a really great hat while shopping for socks. From an artistic perspective, the use of indeterminacy to open the set of considered possibilities corresponds to methods employed by postmodern practitioners such as John Cage [7] and Marcel Duchamp [25]. This corresponds to the way memory and cognition work when the brain is in an intuitive state. "Flat activation" of a greater set of memories that participate in a cognitive moment increases the potential for new associations to emerge [12]. Cognitive models of creativity also utilize indeterminacy [10]. CollageMachine's generative approach to browsing extends experience beyond the routine to expand conceptual spaces and support brainstorming.

Associative Model

The associative model consists of the content elements and multiple referential structures that link them. This model is currently maintained only for the duration of a single session. Each content element includes a tuple of attributes. Information elements inherit attributes from their containers. Some of these attributes represent properties of the content. The *generation* attribute counts the number of links traversed from the session's origin; it is used to effect breadth-first, rather than depth-first, web space retrieval. The *numLinks* attribute counts hyperlinks to the element; as in the Google PageRank metric, [30] more links to a single content element means more importance. The mime attribute is used to give JPEG images priority, since they are usually photographs, and make for interesting recombinant compositions.

Other attributes represent properties of the user and the visual composition designer. *UserSignificance* can be positive or negative. It represents the intensity of the user's interest or disinterest with regard to the element. This field gets modified both by expressive direct manipulation of a visualized media element in the information space, and by inference. We also measure positive clicks, in order to keep track of direct manipulations separately from inference mechanisms. This enables CollageMachine to pay particular attention to the participant's design decisions.

An array structure for available containers – those that have not yet been downloaded – and one for available media elements is maintained for efficient weighted random select operations. Hashtables, with URLs as keys, allow for fast lookup to determine, during parsing of a new document, whether a referenced content element has been previously downloaded. The hyperlink structure of containers and media elements allows expressive interaction to be propagated by spreading activation [34] [31]. The links function as edges; the weights measure the flow capacity of the edges. Each interactive operation specifies an activation (positive values) or inhibition (negative values) energy multiplier for its particular flow. When a media element is altered, a lesser alteration is propagated to its associated container, and, if there is one, hyperlink. When a container's user significance is altered, this is propagated directly to the contained media elements.. Such procedures are invoked recursively, with a damping factor of 5/9. (The efficacy of this constant is based on adhoc tuning, tempered by evaluations.) Thus spreading dies out with a sort of half-life.

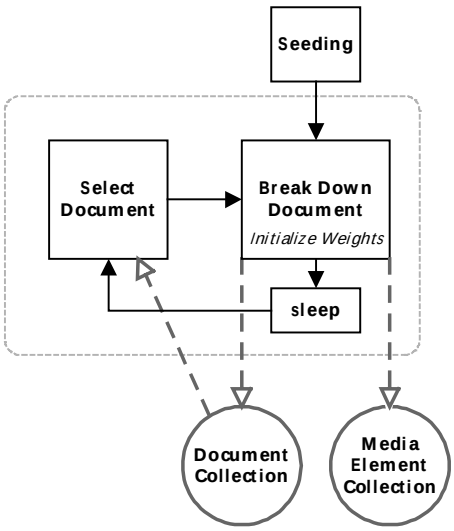


Figure 4. The flow of seeding and breaking down documents.

5. EXPERIENCE AND OPERATIONS

Seeding

The recombinant information space session begins with the specification of a set of initial web addresses, or *seeds* (See Figure 4.). These are the first documents that are fed to the collaging engine. They can be static or dynamic documents; all that matters to the program is that they are in the form of

HTML, and that they can be accessed over the network via the http or ftp protocols, or from a local file system.

The process of seeding is closely related to the context of deployment. As this tool can be used in different ways, the growing set of seeding mechanisms is diverse. Connection of seeding processes and content with activity contexts enables the generation of situated [40] knowledge spaces.

Seeding decisions themselves can be made dynamically, using a database and server-side logic, as in the JumboScope exhibition [22]. In that case, the seeds were samples that served as access points for the visualization of an evolving hypermedia repository, that represented the Tufts University community. This visualization was accessible in a central public space on a flat panel, and also via the web.

Static seeds could be fixed in advance, as part of web site authoring. In this case, CollageMachine functions as alternative metanavigation. This was done by the God Ist Ein DJ project at the Ars Electronica Center. The metanavigation approach would also be an effective means for presenting the contents of any digital library or electronic catalog, in which a substantial investment has been made in the production of high quality images. Art museum and clothing store web sites are examples.

In the standard, user-oriented interface [17], sessions can be seeded via 3 different pathways. In the first, the user types web addresses directly. This extends the paradigm of typing directly into the browser's location/address field; by providing multiple fields, allowing mixed browsing to multiple concurrent destinations. The search engine seeding mechanism works similarly, providing multiple fields that enable the collaging of multiple search engine requests. In this mode, CollageMachine functions as a front-end to Google, providing an alternative form of search results to the standard text listing. It would be interesting to integrate CollageMachine directly with Google. In that case, a persistent form of the CollageMachine associative model could be used by the search engine, as it forms query responses, to augment its understanding of the user's intentions. This would create a new kind of feedback.

The third seeding mechanism, collections, offers a set of pre-curated information space seeds. Popular cultural examples include news (BBC, New York Times, CNN, ABC), art museums (Louvre, Van Gough Museum, British Museum, MOMA, Metropolitan, National Gallery), and 911 (September 11th). Other pre-curated sets have corresponded to events. Ensembles of hypertext literature authors have seeded CollageMachine with their works to produce a dynamic composite information space in public presentation venues, such as the Guggenheim Museum, and the NYU Center for Digital Multimedia. Collections have also been created using the works of peer participants for symposia, such as at the Banff Centre New Media Institute. In these cases of live events, a social space is created by the dynamic reassemblage of digital works in conjunction with a community of participants who are also physically present.

Runtime Parameters

In the standard web incarnation, the user is afforded several other runtime parameter options. Choices for information space size include "full screen," "near full" (which leaves convenient space for the task bar at the bottom of the screen), "almost half," and "quarter" The "stay close" or "allow wandering" choice lets the user focus or limit the web space traversed by a session. If stay close is specified, then the web crawler will be limited to URLs beneath the web address directories specified by the seeds. For example, if the only seed is "www.nytimes.com," then having this string as a prefix is a requirement for traversal of hyperlinks. The set of seeds then defines a spanning space for hyperlink traversals during the course of the session. If an address within that spanning space turns out to refer to a server-side redirect, the redirect target will also be added to the spanning space of traversal. The need for this redirect mechanism was established through experiences with the CNN web site:, where, for example, cnn.com/sports is a server-side redirect to sportsillustrated.cnn.com.

Another runtime parameter, "as found," or "abstracted," maps to two internal parameters. Abstracted allows for the pixelation of smaller images into larger visuals. This creates a textural effect. Abstracted also allows increases the size of the largest possible images. While larger images may look quite interesting visually, they occlude a larger subset of the media elements one could see. When using CollageMachine in order to sift through quantities of elements, this may be considered undesirable.

Breaking Down Documents

When the user pushes the "Launch CollageMachine" button, the seeds are subjected to the same procedures as web pages that are subsequently chosen via the web crawling selection mechanism.[3] Each is downloaded, and parsed. The HTML serves as a guide for how to break the documents down into containers of constituent media elements and references to other containers. The treatment of the **`<img`** tag as a reference to an image element is obvious. Markup, such as the **`<p>`**, **`<div>`**, **`<td>`**, **`<li>`**, and **`<br>`** tags, is used as a first pass in the delimiting of text chunk elements. Text chunks are currently also parsed for sentence boundaries. Their size is thus currently on the order of a sentence, or smaller. Extremely long sentences are currently thrown away. This is one development process aspect with a collaging browser that recurs. Some features are not obliged to work perfectly – only reasonably well. Hyperlinks are translated into container structures, and references thereto, well before downloading. Or, if a referenced document or image is already known to the application, the reference counter attribute of the model, *numLinks*, is incremented.

Figure 5. Temporal controls.

Temporal Controls

The participant can effect temporal development through a subset of the tape recorder and automobile accelerator metaphors. A pause/play button allows the information space evolution to be temporarily stopped, and then resumed. A rate slider affords control over the rate at which documents are downloaded and media elements are added to the visual

[3] After seeding, a thread is started which loops forever. Each iteration selects a container reference from the currently available set, downloads and processes the web page, and sleeps briefly. This is the web crawler.

composition.[4] These controls play an important role in giving the participant control of the experience. In addition, with support for drag and drop, the paused recombinant information space, like Hunter Gatherer, [35] can function as a canvas for collection-building activities.

Building the Recombinant Visualization

Concurrent with the start of breaking down documents, a window appears. The recombinant visualization develops here. The visual composer thread iterates forever through these operations: 1) select a media element from the currently available set to add to the visualization; (2) update the weights for each information space element already onscreen, in case their attributes have changed, and including a history factor that measures how long each has been part of the session;[5] (3) sort the already on screen elements by their weights, so they can be displayed with z-stacking order corresponding to their relative importance;[6] (4) choose a size for the new media element; (5) choose a location on screen for placement of the element; (6) repaint the entire visual composition; and (7) sleep briefly.

Figure 6. CollageMachine Toolbar.

Expressive Interactive Interface

The goal for CollageMachine interactivity is to create an expressive interface, which enables the participant to: take an active role in designing the visual appearance of the information space; and to steer the directions the program takes in traversing web space, over the course of a session. The participant *creates* the experience of receiving information, instead of just passively receiving it in the precise form of someone else's design. In this way, the information space is mutable. The interactive interface maps together the participant's expression of intention with her/his direct manipulation of information space design. This mapping enables the participant to effect what the visual display looks like, and express dis/interest, with a single gesture. The isolated ranking interaction of the typical recommender system, [3] in which the user must perform the ranking task for the system, without immediate gratification, is eliminated by this fusion.

In the current implementation, CollageMachine presents a modal interface, utilizing the MacPaint paradigm, in which a set of tools is made available through a tool palette; one tool (or mode) at a time, is activated. Each tool can be utilized via click, and in some cases, drag and drop operations, with respect to a information space element. The tool acts upon the selected media element, on its container, and, if present, on the target container of a hyperlink. In the case of text elements, which may contain embedded hyperlinks, clicking on or off those links changes the effect. This set of targets forms the operand context for the tool's action. Operation is further propagated from the operand context to related elements by spreading activation, or in the case of negative expression tools, spreading inhibition, to related content elements, based on hyperlink and container relationships.

The *Web Page* tool is active on startup. This tool provides a mapping from recombinant information space browsing to the traditional web browsing paradigm. It opens the target document – either the hyperlinked document if there is one, otherwise the container – in a button-less web browser window. The same window is reused (and brought to the top) for subsequent Web Page tool invocations. This tool is initially active, because usability tests have indicated that this function is the one typically expected by naive users. When this tool is active, clicking works in a manner similar to that of conventional browsers.

The *Positive Grab* and *Negative Grab* tools operate similarly. They enable the user to express interest or disinterest in the operand context of an element. Positive Grab effects "Bring to Top" on click. Both tools enable dragging. Dragging is considered to be the strongest expression. Elements can be dragged within the recombinant space, or beyond its borders, to another program, such as a web browser, content authoring environment, or document editor. The Negative Grab tool was created in response to usability testing [20]. Tests indicated that users sometimes drag elements they like out of the way, so they no longer occlude desirable ones. This is true even though they could use the cut tool to eliminate the undesirable element. An unexpected side effect of this interaction design is that Negative Grab affords drag under. This creates an unusual tactile and visually stimulating experience.

The *Cut* tool's operation is more straightforward. It removes an element from the information space on click. Disinterest is expressed with regard to the tool's operand context.

The Text tool enables editing of existing text elements, and creation of new ones. Through this process, the user is able to annotate the information space, with her/his own conceptual glue. These annotations also enlarge the scope of the ongoing evolution of the space, in that they trigger the automatic synthesis of new Google queries.

Image Processing Techniques

CollageMachine uses image processing techniques to create visual layering. They effect treatments on and fastenings of information elements. These treatments are mapped to serve multiple functions: visual, cognitive, and operational. Blur and desaturation create foreground/background effects that aid the viewer in seeing the recombinant space's multiplicity of elements. As elements are progressively desaturated, reflecting their age on screen, their color seems to be gradually drained out of them. This mirrors the way human memory works; events fade over time. This is similar to Hollan and Hill's notion of edit wear [17]. Similarly, negative interaction effects blur on the elements in the operand context (See Figure 7.); this blur is propagated to "related" elements via spreading inhibition. The

[4] Work in progress will extend these temporal controls to include reverse, and will differentiate play and record.

[5] The history factor is only relative, based on an element's age compared to other elements in the collage. It does not absolutely account for elements that persist due to user significance measures in the model.

[6] Using a form of radix sort, the elements in each fat pixel are sorted as part of the same procedure that sorts the whole collage.

Figure 7. Successive states of disinterest expressed through blur, and propagated through spreading inhibition..

blurred and less colorful elements become less prominent visually, than those that are sharp and colorful. As these elements fall into the background, newer and more important elements grow easier to see. These effects occur automatically, based on the participant's perceptual + cognitive systems. Further, the participant who understands these mappings can use them more intentionally. As these imagistic mappings visualize the state of the program's model, so they inform the participant about how the program is working, and, therefore, about what it is likely to do. Positive interaction with an information element overrides the desaturation function, maintaining an element's visual state and prominence. This is one of the many ways in which CollageMachine works to respect and reinforce the participant's direct design decisions.

Rollover State

A number of fluid visual modifications are effected on rollover of an information element (See Figure 8.). Immediately, the unprocessed version is displayed, removing all effects, including blur, desaturation, and alpha blending. This reminds the participant about what the element looks like in its original context. The element is surrounded by a black selection box. Further, if the focus information element is partially occluded by other elements, the bounding regions of occlusion are drawn with a window shade horizontal hashing. This makes visible the area that the focus element would take in the information element if it was not occluded. Once rollover lasts long enough to be deemed an intentional exploration rather than an accidental or in-passing wave -- for 700 milliseconds – a bubble above the element displays metadata about it's operand context. For example, in what document is it contained? To what document is it perhaps hyperlinked? If it is an image, is there metadata in the HTML markup, via the `alt` text attribute? The participant does not need to invoke the web page tool to access this contextual information.

6. FUTURE WORK

The range of significant means for improvement, and applications of CollageMachine is surprisingly vast. We are currently working to extend the associative model of content to incorporate information retrieval's term vector model [34]. We also plan to extend the model of the participant beyond a simple measure of interest, to account for a range of feelings, interests, physical sensations, cognitive processes, and symbolic senses. An expanded set of interactive design capabilities will give the participant further ability both to effect the visualization and to express intention. We are beginning to work on integrating a computer vision system that will enable 3-D physical interaction via gesture. The relationship between physicality and expressivity will be explored. Ubiquitous displays and situated integration will relate the generation of digressive information brainstorm spaces more directly and automatically to participants' ongoing activities. Applications areas such as digital libraries, authoring, and local and remote collaboration will also be addressed.

7. ACKNOWLEDGMENTS

Research has been supported by Texas A&M Department of Computer Science and Center for Digital Libraries, TEES, Creating Media, NYU Media Research Lab, NYU Department of Computer Science, and NSF Grant GER-9454173. Thanks to John Leggett, Rick Furuta, Frank Shipman, Ken Perlin, Richard Schechner, Barbara Kirshenblatt-Gimblett, and Thecla Schiphorst. Dominic Girard, Eli Levin-Goldstein, and Jacques Lowe have contributed to the visual/interactive design elements.

8. REFERENCES

[1] Adamowicz, E. *Surrealist Collage In Text And Image.* Cambridge University Press. 1998.

[2] Anderson, C. and Horvitz, E. Web Montage: A Dynamic Personalized Start Page. *Proc WWW 2002*.

[3] Aragon, Louis. *Les Collages*. Paris: Hermann, 1965.

[4] Balabanovic, M. *Learning to Surf: Multiagent Systems for Adaptive Web Page Recommendation.* http://www.rsv.ricoh.com/~marko/.

[5] Bender, W. Thoughts on the future of news. http://web.media.mit.edu/~walter/

[6] Bush, V. As We May Think. Atlantic Monthly, July, 1945.

[7] Cage, J. *Silence*. Middletown CT, Wesleyan Univ Press. 1961.

[8] Chesnais, P. et al. The Fishwrap Personalized News System. Proc 2nd IEEE Intl Workshop on Community Networking (1995), 275-282,

[9] Duda, R., Hart, P. and Stork, D., Pattern Classification, Wiley, New York, 2001.

[10] Finke, R., Ward, T., and Smith, S. *Creative Cognition*. Cambridge MA., MIT Press, 1992.

[11] Furuta, R. Shipman, F., Marshall, C. Brenner, D., Hsieh, H.. Hypertext paths and the World-Wide Web: experiences with Walden's Paths. *Proc ACM Hypertext*, 167-176, 1997.

[12] Gabora, L. & Aerts, D. Contextualizing concepts using a mathematical generalization of the quantum formalism. *Journal of Experimental and Theoretical Artificial Intelligence*. (in press)

[13] Gruhl, D. and Bender, W. A New Structure for News Editing. *IBM Systems Journal* 39: 3&4: 569-588, 2000.

[14] Hill, W., and Hollan, J.. "History-Enriched Digital Objects", Proceedings of the ACM Conference on Computers, Freedom and Privacy , 1993.

[15] Hopps, W. and Davidson, S. Robert Rauschenberg, a Retrospective. New York: Guggenheim Museum, 1997.

[16] Karadkar, U.P., Kerne, A., Furuta, R., Francisco-Revilla, L., Shipman, F., Wang, J., Connecting Interface Metaphors to Support Creation of Hypermedia Collections, Proc European Conference on Digital Libraries, in press, 2003.

[17] Kerne, A. *CollageMachine*. http://csdl.tamu.edu/collagemachine

[18] Kerne, A. *CollageMachine: A Model of Interface Ecology*. NYU Ph.D. Dissertation. 2001.

[19] Kerne, A. CollageMachine: An Interactive Agent of Web Recombination. *Leonardo* 3:5, 347-350, Nov 2000.

[20] Kerne, A. CollageMachine: Interest-Driven Browsing Through Streaming Collage, *Proc Cast01: Living in Mixed Realities*, Bonn, Sept 2001.

[21] Kerne, A. CollageMachine: Temporality and Indeterminacy in Media Browsing via Interface Ecology, in *Proc CHI 97- Extended Abstracts*, 297-298, 1997.

[22] Kerne, A. Concept-Context-Design: A Creative Model for the Development of Interactivity", *Proc ACM Creativity and Cognition 4*, 2002.

[23] Krauss, R. *The Optical Unconscious*. Cambridge, MIT Press, 1998.

[24] Lieberman, H. Letizia: An Agent That Assists Web Browsing. *Intl Joint Conference on Artificial Intelligence*, Montreal, August 1995.

[25] Lippard, L. *Dadas on Art*. Englewd Cliffs NJ, 1971.

[26] Lippard, L. *Six Years: The dematerialization of the art object*, Los Angeles, UCLA Press, 1973.

[27] Manovich, L. The Language of New Media. Cambridge: MIT Press, 2001.

[28] Miles, A. Hypertext Structure as the Event of Connection. Proc Hypertext 01, Bergen, 2001.

[29] Nelson, T. Computer Lib, 1974.

[30] Page, L., Brin, S., Motwani, R., Winograd, T. The PageRank Citation Ranking: Bringing Order to the Web. http://dbpubs.stanford.edu:8090/pub/1999-66. 1998.

[31] Pirolli, P., Pitkow, J., and Rao, R.. "Silk from a Sow's Ear: Extracting Usable Structures from the Web", Proc CHI '96 Vancouver, 1996.

[32] Public Enemy. *It takes a nation of millions to hold us back*. New York, Def Jam, 1995.

[33] Sadie, S., ed. *New Grove Concise Dictionary of Music*. Macmillan: Hampshire, UK, 1994.

[34] Salton, G. and Buckley, C. On the Use of Spreading Activation Methods in Automatic Information Retrieval. *Proc SIGIR 1988*. 147-160.

[35] Schraefel, M. et al. Hunter Gatherer: Interaction Support for the Creation and Management of Within-Web-Page Collections. *Proc WWW 2002*.

[36] Sherman, C. FAST Sprints Past Google in Search Engine Size Wars. SearchDay 291, June 17, 2002. http://searchenginewatch.com/searchday/02/sd0617-update.html

[37] Shipman, F., Hsieh, H., Airhart, R., Maloor, P. and Moore, J.M., The Visual Knowledge Builder: A Second Generation Spatial Hypertext, Proc ACM Conference on Hypertext, 113-122, 2001.

[38] Spies, W. *Max Ernst Collages: The Invention of the Surrealist Universe*. NY, Harry Abrams. 1988.

[39] Stockhausen, K. Elektronische Musik. 1952-1960. Germany, Kurten. 1991.

[40] Suchman, L. *Plans and Situated Actions*. Cambridge UK, Cambridge University Press. 1987.

[41] Wardrip-Fruin, N. Impermanence Agent. http://www.cat.nyu.edu/agent

[42] Wisniewski, M. Netomat. http://www.netomat.net

Expressive AI: A Semiotic Analysis of Machinic Affordances

Michael Mateas
Georgia Institute of Technology
Literature, Communication and Culture & The College of Computing
686 Cherry Street
Atlanta GA 30032-0615
michael.mateas@lcc.gatech.edu

ABSTRACT
Expressive AI is a hybrid practice, combining artificial intelligence (AI) research and art making, that simultaneously focuses on the negotiation of meaning mediated by an art object *and* the internal structure of AI systems. These two apparently disparate views are unified through the concept of affordance: negotiation of meaning is conditioned by interpretive affordances while the internal structure of the AI system is conditioned by authorial affordances. This paper employs a structuralist semiotic analysis to unpack the notion of interpretive and authorial affordance, exploring the deep relationships between AI code structures, authorial intentionality, and culturally negotiated meaning.

Keywords
Artificial intelligence, semiotics, art

1. Introduction
Art and artificial intelligence (AI) research appear to be quite different practices. Where art practice focuses on the negotiation of meaning as mediated by the art object, AI research focuses on internal system structure and the interaction between system and environment. My work in AI-based art and entertainment simultaneously engages in AI research and art making, a research agenda and art practice I call *Expressive AI* [10, 11].

Expressive AI has two major, interrelated thrusts: (1) *exploring the expressive possibilities of AI architectures* – posing and answering AI research questions that wouldn't be raised unless doing AI research in the context of art practice, and (2) *pushing the boundaries of the conceivable and possible in art* – creating artwork that would be impossible to conceive of or build unless making art in the context of an AI research practice.

Expressive AI is thus a hybrid practice simultaneously focusing on the negotiation of meaning *and* the internal structure of AI systems. These two apparently disparate views are unified through the concept of affordance: negotiation of meaning is conditioned by interpretive affordances while the internal structure of the AI system is conditioned by authorial affordances. In [11] I described how a focus on authorial expression changes the AI research agenda, positioned Expressive AI relative to both symbolic and embodied AI, and introduced the idea of interpretive and authorial affordance. This paper employs a structuralist semiotic analysis to unpack the notion of interpretive and authorial affordance, exploring the deep relationships between AI code structures, authorial intentionality, and culturally negotiated meaning.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

2. Example Systems
This section provides brief descriptions of three AI-based artworks. These systems are used as examples throughout the rest of the paper.

2.1 Office Plant #1
Walk into a typical, high tech office environment, and, among the snaking network wires, glowing monitors, and clicking keyboards, you are likely to see a plant. In this cyborg environment, the silent presence of the plant fills an emotional niche. Unfortunately, this plant is often dying; it is not adapted to the fluorescent lighting, lack of water, and climate controlled air of the office. *Office Plant #1* [5] is an exploration of a technological object, adapted to the office ecology, that fills the same social and emotional niche as a plant. *Office Plant #1* employs text classification techniques to monitor its owner's email activity. Its robotic body, reminiscent of a plant in form, responds in slow, rhythmic movements to express a mood generated by the monitored activity. In addition, low, quiet, ambient sound is generated; the combination of slow movement and ambient sound thus produces a sense of presence, responsive to the changing activity of the office environment.

Figure 1. Office Plant #1.

Office Plant #1 classifies incoming email into social and emotional categories using AI statistical text classification techniques. Given the categories detected by the email classifiers,

a Fuzzy Cognitive Map (FCM) determines which behavior the plant should perform. The FCM is a neural network-like structure in which nodes, corresponding to behaviors, are connected to each other by negative and positive feedback loops.

2.2 Terminal Time

Terminal Time [14] is a story generation system that constructs ideologically-biased documentary histories, consisting of spoken narrative, video sequence and sound track, in response to audience feedback measured by an applause meter. One of the goals of *Terminal Time* is to build a caricature model of the documentary film production process. Rather than "objectively" reporting a sequence of events through the eye of a camera (the implied production process in documentary film), events are instead selected and biased so as to satisfy an ideological position, assembled into a desired narrative, and only then is video footage selected to illustrate the constructed narrative. As a large-audience interactive artwork, *Terminal Time* allows an audience to explore the role of ideological bias in the construction of history. As an AI research system, *Terminal Time* integrates a novel model of ideologically-biased reasoning within a story-generation framework.

The architecture makes use of several representations and knowledge sources including: a knowledge base of historical events represented in an ontology based on the Upper Cyc Ontology, ideologue-specific representations of rhetorical goals that select and "spin" events, rhetorical devices that can be used to "glue" spins together to form historical narratives, a plan-based natural language generator, and a database of term-indexed video clips.

2.3 Façade

Façade is an artificial intelligence-based art/research experiment in electronic narrative – an attempt to move beyond traditional branching or hyper-linked narrative to create a fully-realized, one-act interactive drama [12, 13]. *Façade* incorporates the player's interaction with autonomous characters into a well-shaped dramatic arc with a clear inciting incident, progressive complication leading to a climax, and closure. In *Façade*, you, the player, play the character of a longtime friend of Grace and Trip, an attractive and materially successful couple in their early thirties. During an evening get-together at their apartment that quickly turns ugly, you become entangled in the high-conflict dissolution of Grace and Trip's marriage.

Architecturally, *Façade* consists of a number of components. ABL (A Behavior Language) is a novel reactive planning language for authoring believable agents. ABL provides language support for authoring coordinated, multi-character dramatic action. The drama manager operationalizes dramatic beats. In dramatic writing, a beat is the smallest unit of dramatic value change, where dramatic values are properties of individuals or relationships such as trust, love, hope, etc. In *Façade* beats are architectural entities, consisting of preconditions, a description of the values changed by the beat, success and failure conditions, and joint behaviors (written in ABL) that coordinate the characters in order to carry out the specific beat. The drama manager attempts to sequence beats so as to incorporate player interaction while making specific dramatic arcs (value change graphs) happen. The natural language processing system employs semantic parsing to map dialog typed by the player into discourse acts (e.g. agree, disagree) and interprets the resulting discourse acts as a function of the current discourse context (most often defined by the currently active beat). Finally, a custom non-photorealistic animation engine presents the story world as a real-time, 3D space through which the player can move, gesture, interact with objects, and talk with characters (dialog input is accomplished through typing).

3. Affordances

The notion of affordance was first suggested by Gibson [8] in his theory of perception and was later re-articulated by Norman [17] in the field of interface design. For Gibson, affordances are *objective*, actionable properties of objects in the world. For an animal to make use of the affordance, it must of course perceive it in some way, but for Gibson, the affordance is there whether the animal perceives it or not; an unperceived affordance is waiting to be discovered. For Norman, affordances become *perceived and culturally dependent*. That is, rather than viewing the relationship between sensory object and action as an independent property of the object+animal system, this relationship is contingent, dependent on the experiences of the perceiver within some cultural framework. For example, for a person who has spent the last 10 years using the web, blue underlined text now affords an action, clicking with a pointing device, with the expectation that this clicking will "follow a link" to another information node. If blue underlined text is used in a different interface merely as a way to emphasize text, this is likely to generate confusion because the hypothetical interface is violating an affordance. It is this second notion of contingent affordance that I use here. But note that though affordances are contingent, they are not arbitrary – affordances are conditioned by the details of human physiology (what we can sense, how our bodies move), by cultural memory, and by the perceivable physical properties of objects. While new affordances can come into existence, as illustrated by the link-following affordance of blue underlined text, these innovations are conditioned by earlier affordances (e.g. the physical affordances of computer mice) and take active cultural work to establish.

3.1 Interpretive Affordance

Interpretive affordances support the interpretations an audience makes about the operations of an AI system, conditioning the meanings negotiated between artist and audience. Interpretive affordances provide resources both for narrating the operation of the system, and additionally, in the case of an *interactive* system, for supporting intentions for action.

For AI-based art, narrative affordances support the audience in creating a story about the operation of the piece and how this operation relates to the artist's intention. For example, imagine having *Office Plant #1* on your desk. The name, plus the physical form, prepares one to view the sculpture as a plant – it has identifiable parts that metaphorically relate to the stem, flower, and leafs of biological plants. The wooden box of the base, hammered finish of the flower, and whimsical piano-wire fronds topped with crinkled, copper-foil-wrapped spheres, give the plant a non-designerly, hand-built look that communicates that it is neither a consumer electronic toy nor serves any functional purpose. Yet it is clearly a machine – it hums quietly while operating, moves very slowly (the motion is visible only if you watch patiently), and, when returning to the desk after an absence,

is sometimes in a different configuration than it was left in. The plant starts moving when email is received; over time one can notice a correlation between the plant's physical poses and the email received. All of the perceived features of the plant, the materials used and the details of fabrication, the physical form, the temporal behavior, the relationship between this behavior and email, constitute the narrative affordances, the "hooks" that the plant's owner uses to make sense of the plant, to understand the plant in relationship to themselves and their daily activity.

For interactive art, intentional affordances support the goals an audience can form with respect to the artwork. The audience should be able to take an action and understand how the artwork is responding to this action. This doesn't mean that the artwork must provide simple one-to-one responses to the audience's actions. Such simple one-to-one responses would be uninteresting; rather, the poetics of the piece will most likely avoid commonly used tropes while exploring ambiguities, surprise, and mystery. But the audience should be able to understand that the system is responding to them, even if the response is unexpected or ambiguous. The audience should be able to tell some kind of unfolding story about their interaction with the work. Both the extremes of simple stereotyped responses to audience interaction making use of well-known tropes, and opaque incoherence with no determinable relationship between interaction and the response of the art work, should be avoided.

A concern with interpretive affordances is often alien to AI research practice. Though the role of interpretation is sometimes discussed (e.g. the Turing test is fundamentally about interpretation [20], Newell's knowledge level is an attribution made from *outside* an AI system [15]), most often AI systems are discussed in terms of *intrinsic* properties. But for artists, a concern with interpretive affordance is quite familiar; negotiating meaning between artist and audience is central to artistic practice. Expressive AI adopts this concern within the context of AI-based art. But Expressive AI also adopts a concern for the internal functioning of the artifact from AI research practice.

3.2 Authorial Affordance

The authorial affordances of an AI architecture are the "hooks" that an architecture provides for an artist to inscribe their authorial intention in the machine. Different AI architectures provide different relationships between authorial control and the combinatorial possibilities offered by computation. Expressive AI engages in a sustained inquiry into these authorial affordances, crafting specific architectures that afford appropriate authorial control for specific artworks.

This concern with the machine itself will be familiar to AI research practitioners. However, AI research practice often downplays the role of human authorship, focusing on the properties of the architecture itself independent of any "content" authored *within* the architecture. Multiple architectures are most often compared in a content-free manner, comparing them along dimensions and constraints established by theories of mind, or theories of brain function (not necessarily at the lowest, neuron level), or comparing their performance on established benchmark problems. For Expressive AI, the concern is with how the internal structure of the machine mediates between authorship and the runtime performance.

A focus on the internals of the machine itself is often alien to current electronic media practice; the internal structure of the machine is generally marginalized. The machine itself is considered a hack, an accidental byproduct of the artist's engagement with the concept of the piece.

> One might generalize in this way (with apologies to both groups): artists will kluge together any kind of mess of technology behind the scenes because the coherence of the experience of the user is their first priority. Scientists wish for formal elegance at an abstract level and do not emphasize, or do not have the training to be conscious of inconsistencies in, the representational schemes of the interface. [18]

In discussions of electronic media work, the internal structure of the machine is almost systematically effaced. When the structure is discussed, it is usually described at only the highest-level, using hype-ridden terminology and wishful component naming (e.g. "meaning generator", "emotion detector"). At its best, such discursive practice is a spoof of similar practice within AI research, and may also provide part of the context within which the artist wishes her work to be interpreted. At its worst, such practice is a form of obfuscation, perhaps masking a gap between intention and accomplishment, the fact that the machine does not actually do what is indicated in the concept of the piece.

Yet it is nonetheless the case that an artist's concern with the coherence of the audience experience, with the crafting of interpretive affordances, is entirely appropriate – creating an audience experience is one of the primary reasons the artwork is being made in the first place. So why should an artist concern herself with authorial affordances, with the structural properties of the machine itself? Because such a concern allows an artist to explore expressive possibilities that can only be opened by a simultaneous inquiry into interpretive affordance and the structural possibilities of the machine. Interpretive and authorial affordances are coupled – a concern with the machine enables audience experiences that aren't achievable otherwise.

3.3 Combining Interpretive and Architectural Concerns

The splitting of AI-based art practice into interpretive and authorial concerns is for heuristic purposes only, as a way to understand how Expressive AI adopts concerns from both art practice and AI research practice. Expressive AI practice combines these two concerns into a dialectically related whole; the concerns mutually inform each other. The "interface" is not separated from the "architecture". In a process of total design, a tight relationship is maintained between the sensory experience of the audience and the architecture of the system. The architecture is crafted in such a way as to enable just those authorial affordances that allow the artist to manipulate the interpretive affordances dictated by the concept of the piece. At the same time, the architectural explorations suggest new ways to manipulate the interpretive affordances, thus suggesting new conceptual opportunities. Thus both the artist's engagement with the inner workings of the architecture and the audience's experience with the finished artwork are central, interrelated concerns for Expressive AI.

The AI-based artist should avoid architectural elaborations that are not visible to the audience. However, this admonition should not be read too narrowly. The architecture itself may be part of

the concept of the piece, part of the larger interpretive context of people theorizing about the piece. For example, one can imagine building a machine like *Terminal Time* in which some small collection of historical narratives have been prewritten. The narrative played is determined by a hard-coded selection mechanism keyed off the audience polls. For any one audience, the sensory experience of this piece would be indistinguishable from *Terminal Time*. However, at a conceptual level, this piece would be much weaker than *Terminal Time*. A *Terminal Time* audience is manipulating a *procedural process* that is a caricature of ideological bias and of institutionalized documentary filmmaking. The operationalization of ideology is critical to the concept of the piece, both for audiences and for artists and critics who wish to theorize the piece.

4. The Code Machine and the Rhetorical Machine

AI (and its sister discipline Artificial Life), consists of both technical strategies for the design and implementation of computational systems, and a pared, inseparable, tightly entangled collection of rhetorical and narrative strategies for talking about and thus understanding these computational systems as intelligent, and/or alive.

These rhetorical strategies enable researchers to use language such as "goal", "plan", "decision", "knowledge", to simultaneously refer to specific computational entities (pieces of program text, data items, algorithms) and make use of the systems of meaning these words have when applied to human beings. This double use of language embeds technological systems in broader systems of meaning.

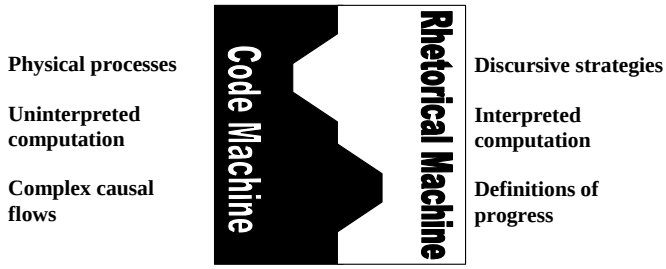


Figure 2. Total system = code machine + rhetorical machine

There is an uncomfortable relationship between a purely relational (and thus literally meaningless) technical manipulation of computational material, and the interpretation of this computational material by a human observer. Simon and Newell posited the physical symbol system hypothesis as a fundamental assumption of AI [16]. This hypothesis states that a physical system consisting of a material base that can take on various configurations (call these configurations "symbols") and a material process that manipulates these physical constellations to yield new constellations is sufficient for the production of intelligent behavior. This formulation immediately produces an interpretation problem in which an external observer is necessary in order to view the material constellations as signs in such a manner that intelligence can be observed in the material production of sign from sign. Interpretation, with all of its productive open-endedness, is thus crucial to the definition of intelligent system, but is usually pushed to the background of AI practice.

The necessity of rhetorical strategies of interpretation is not avoided by "subsymbolic" techniques such as neural networks or genetic algorithms utilizing numeric genomes (i.e. not the tree-shaped, symbolic genomes of genetic programming), nor by machine learning methods based on generalization from training data, nor by behaviorist robotic techniques that link sensors to effectors through stateless combinational circuitry or finite state machines. These approaches still require the interpretation of an observer in order to make sense of the input/output relationships exhibited by the system, to select the primitive categories (features) with which the inputs are structured, and to tell stories about the processes producing the input/output relationships. These stories are essential for thinking through which technical constructions to try next, that is, for simultaneously defining a notion of progress and a collection of incremental technical constructions that make progress according to this notion.

The rhetorical strategies used to narrate the operation of an AI system varies depending on the technical approach, precisely because these interpretative strategies are inextricably part of the approach. Every system is doubled, consisting of both a computational and rhetorical machine (see figure 2). Doubled machines can be understood as the interaction of (at least) two sign systems, the sign system of the code, and a sign system used to interpret and talk about the code.

The central problem of AI is often cast as the "knowledge representation" problem. This is precisely the problem of defining structures and processes that are *simultaneously* amenable to the uninterpreted manipulations of computational systems *and* to serving as signs for human subjects. This quest has driven AI to be the most promiscuous field of computer science, engaging in unexpected and ingenious couplings with numerous fields including psychology, anthropology, linguistics, physics, biology (both molecular and macro), ethnography, ethology, mathematics, logic, etc. This rich history of simultaneous computational and interpretive practice serves as a conceptual resource for the AI-based artist.

The relationship between the sign system of the code (the code machine) and the sign system used to talk about the code (the rhetorical machine) can be explicated via a semiological analysis. By semiology, I mean the semiotic tradition following Saussure's General Linguistics [19], and explicated by thinkers such as [9, 4]. The treatment in this paper most closely follows Barthes [3, 4].

4.1 The Code System

The program code, considered as a sign system, relates two planes: a plane of expression containing the space of all possible pieces of program text (the marks on a screen or page), and a plane of content containing the space of all potential executions. That is, a piece of program code is a signifier signifying (the mental concept of) the effect of executing this code. For example, the signified of the simple sign (code fragment) `x = 1` is, for programmers used to working in imperative languages, probably something like placing a 1 within a box labeled x.

Note that code signs, as is the case with any sign, provide no privileged access to an unmediated reality. The signified is the mental concept of an execution, not the execution itself. The relationship between the mental concept of an execution and the physical effect of executing a piece of code on a concrete

computer (e.g. for contemporary digital computers, changing voltage levels in pieces of silicon) falls outside of the purview of structuralist semiotics. A code fragment is a sign-function, having both a utilitarian, technical use (the physical effect of executing the code on a concrete machine), while serving as a sign for its potential execution. Obviously there are constraints imposed on sign value by use value; for example, the physicality of a rubber ball, and the technical functions (e.g. bouncing) that the physicality of a rubber ball supports, prevents (or at least makes quite difficult) the rubber ball from taking on the sign value of a tasty snack. Similarly, the possible sign values of a code fragment are constrained by the use value, the physical effect of its execution on concrete machinery. Though a structuralist semiotic analysis has its limits, such as difficulty in offering a detailed analysis of the relationships between sign and use value, it remains the case that much of human activity is structured by language-like interactions, from which a semiotic analysis gains its traction. In the specific case of the activity of programming, programmers think about potential executions and read and write texts to express those potential executions; this language-like activity suggests that the semiotic view of program code as a sign system, while not explaining *everything* about the human activity of programming, is likely to yield dividends.

To further unpack the idea of code as a semiotic system, consider the example of rhetorical goals in *Terminal Time*. The textual representation, the code, for a specific rhetorical goal appears in Figure 3.

```
(def-rhetgoal
  :name :give-positive-example-of-big-science
  :app-test
   (%and
    ($isa ?event %SciTechInnovationEvent)
    ($performedBy ?event ?bigsci)
    ($isa ?bigsci $LegalGovernmentOrganization)
    ($isa ?bigsci $ResearchOrganization))
  :rhet-plans (:describe-event)
  :emotional-tone :happy)
```

Figure 3. The code representation of a rhetorical goal.

This complex sign is itself a syntagm, composed of a constellation of signs. But considering the complex sign as a unity, the rhetorical goal signifies potential executions in which the system will tend to include a certain class of historical events in the constructed documentary, in this case, events in which governmental research organizations engage in scientific or technical research, in such a way as to make a certain point, in this case, that it is beneficial when science and government come together. It is interesting, perhaps surprising, that this relatively small textual signifier signifies potential executions that relate so directly to *Terminal Time's* output; watching a generated documentary (in which this goal is active) with this code sign in hand, it is possible to relate the appearance of specific historical events in the documentary (such as a breathless, glowing description of the moon landing or the invention of the atomic bomb) to this code sign, that is, to the effect on execution of this textual signifier. It is certainly not a given that a system of code signs would necessarily provide form to the plane of textual representations (expression) and the plane of potential executions (content) in this way. It takes work to articulate the planes in this particular way – this work is in fact the *creation of a custom code system*.

Standard languages, such as C++, lisp, or Java, define code systems, specific ways of chopping up the spaces of textual representations and potential executions. Like many sign-function systems, the more radical innovation of the creation of the sign system lies with special individuals or organizations who define the language, with consumers of the language limited to working with the signs, the associations between text and execution, established by the language. But it is standard practice in computer science, enabled by Turing equivalence, to use a pre-given code system (language) to implement new code systems that provide *different associations between text and execution*. This practice allows individuals to engage in the more radical innovation of creating new code systems particularly suited for a specific task. Mainstream languages, such as the three mentioned above, tend to be strongly procedural; the control structure, which determines the temporal relationship between bits of execution, is explicitly captured in the textual representation. However, this is not the only kind of code system. One can define purely declarative code systems, such as the rhetorical goal above. In declarative systems, the textual representation does not explicitly capture temporal relations in execution. Rather, the code signs indicate execution propensities. The system as a whole will tend to behave in certain ways if the declarative sign is part of the system, though the precise execution path (temporal sequence of sign execution) is unknown. Or the custom language may be a hybrid, such as ABL, which combines the declarative features of production systems with the procedural features of more mainstream languages.

The *architecture* is the conglomeration of code that implements a custom language, that is, establishes the relationship between bits of textual representation and potential executions. For example, in *Terminal Time* a rhetorical goal becomes a sign by virtue of its role within the entire architecture. The rhetorical goal has relationships with or participates in many parts of the architecture, including the knowledge base, the story board (where narrative construction takes place), natural language generation, the selection of music, and (indirectly, through the goal's effect on the natural language generator) the sequencing of video clips. This little bit of text gains its meaning through its effect on a broad array of processes throughout the architecture.

At this point it is possible to provide a semiotic account of the code system properties that yield interpretive and authorial affordances.

4.1.1 Affordance in the Code System

An AI-based artwork is a semiotic system productive of a (potentially large) number of syntagms. AI-based artworks are thus *generative*; computational processes provide the combinatoric machinery necessary to select terms out of the fields of potential terms (associative fields) provided by the system. The system produces variable syntagms in different situations. For example, *Office Plant #1's* behavior over time depends on the email received by its owner, the content of documentaries generated by *Terminal Time* depends on audience answers to the psycho-graphic polling questions, and Trip and Grace's moment-by-moment behavior in *Façade*, as well as the more global story structure, depend on the player's real-time interaction and patterns of interaction over time.

The internal structure of the machine, the program code, wires, circuits and motors out of which a work might be constructed, is

itself a syntagm of the semiotic system defined by the architecture (see Figure 4). The architecture consists of the custom code systems, processes, modules, and relationships between modules, which together define the implementation language, the sign system within which the work will be constructed. Building an AI-based artwork thus means constructing a semiotic system of implementation (an architecture, $system_1$) such that it supports the construction of a syntagm (the specific work built within the architecture, $syntagm_1$), which, when executed, becomes a semiotic system ($system_2$) autonomously productive of its own syntagms ($syntagm_2$) in different situations. $System_1$ (the architecture) has appropriate *authorial affordances* when there is a "natural" relationship between changes to the $syntagm_1$ and changes in the syntagmatic productivity of $system_2$. By "natural" is meant that it is easy to explore the space of syntagmatic productivity consistent with the artistic intention of the piece.

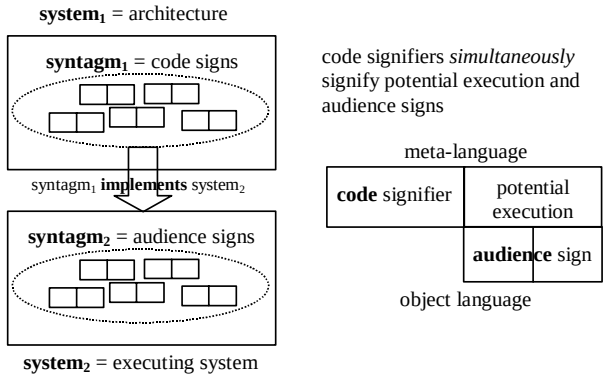


Figure 4. Relationships in the code system.

For example, in *Terminal Time*, the AI architecture is $system_1$. $Syntagm_1$ is the collection of historical events (collections of higher-order predicate calculus statements), rhetorical goals, rhetorical devices, natural language generation rules, rhetorical plans, and annotated video and audio clips, which collectively make up the specific artwork that is *Terminal Time*[1]. Individual signs within $syntagm_1$, as well as $syntagm_1$ as a whole, *are* signs (have meaning) by virtue of their participation within $system_1$. The execution of $syntagm_1$ results in $system_2$, in a runtime instance of *Terminal Time*. And, as the audience interacts with $system_2$, it produces $syntagm_2$, a particular documentary out of the space of all possible documentaries expressible within (producible by) $system_2$. While the structure of $syntagm_2$ is quite literally determined by $system_2$, for the audience, the meanings expressed by $syntagm_2$ are determined by a meshwork of different sign systems, including the system of documentary imagery, the system of cinematic music, the linguistic system for English (the voiceover), and a folk psychology of the execution of $system_2$ (e.g. "we voted that religion is a problem in the world, and now it's trying to make the point that religion is bad"). Thus $syntagm_2$ is multi-articulated; its meaning is determined not just by $system_2$, but also by a number of sign systems *outside* the technical $system_2$.

[1] Since signs may be added or changed over time, such as the modification or addition of rhetorical devices or historical events, *Terminal Time* as a specific piece changes over time.

$System_1$ is a meta-language for talking about $system_2$; utterances in $system_1$ ($syntagm_1$ or fragments) talk about potential utterances of $system_2$ ($syntagm_2$ or fragments) (see Figure 4). For *Terminal Time*, $system_1$ utterances, such as the rhetorical goal in Figure 3, are a way of talking about potential $system_2$ utterances, such as a breathless, glowing description of the invention of the atomic bomb. $System_1$ offers *effective authorial affordances when one and the same* $syntagm_1$ *simultaneously talks about desired* $syntagms_2$ *(or fragments), and, when executed, implements the appropriate* $system_2$ *that indeed produces the desired* $syntagms_2$. This property is not trivial – there are a number of ways in which it can fail to hold.

It can be the case that $system_1$ fails to provide appropriate signs for talking about desired properties of $syntagm_2$. For example, an early version of *Terminal Time's* architecture represented historical events directly at the natural language generation and video clip sequencing level. There was a fairly direct connection between answers to the audience polls and the generation of specific text about specific events. Given this $system_1$, it was impossible to express *general* relationships between poll answers and categories of events. For example, if the winning answer to the question "What is the biggest problem in the world today" is "It's getting harder to earn a living and support a family", the desired $syntagm_2$ should include events demonstrating the evils of capitalism. Given a relatively direct connection between poll answers and natural language generation, there just was no way of expressing this more general desired property of $syntagm_2$, and thus certainly no way of implementing the appropriate $system_2$ with $syntagm_1$.

It can be the case that $syntagm_1$ utterances *purport* to talk about desired $syntagms_2$, but in fact, when executed, don't implement a $system_2$ that produces the desired $syntagm_2$. For example, in *Office Plant #1*, statistical text classifiers map incoming email into social and emotional categories. The categories appearing in an email stream then condition the physical behavior of the device. However, if the email categories are being inappropriately assigned to individual emails, then the decision making process that uses the assigned categories to decide which physical behaviors to perform will make inappropriate decisions. That is, the author will think that they're specifying a $system_2$ that reacts in a specific way to, for example, an apology email, when in fact the internal label apology (a sign in $syntagm_1$) does not properly correspond with the intuitive notion of an apology. Thus the statistical text classifiers must be trained in such a way that the labels (categories) produced by the classifiers have an appropriate correspondence with email messages.

As a final example of the failure of authorial affordance, it can be that case that $syntagm_1$ is successful in simultaneously describing a desired $syntagm_2$ and implementing an appropriate $system_2$, but that, when the audience (who may in fact be the same as the author) actually experiences the produced $syntagm_2$, its interpretation is different than expected. This situation arises precisely because $syntagm_2$ doesn't participate in just the technical $system_2$, but in a meshwork of sign systems *outside* of the technical system. That is, part (perhaps a large part) of the meaning of $syntagm_2$ is opaque to the technical system, but rather comes along for the ride as the technical system manipulates and produces signs. For example, in *Façade*, a beat, and the associated beat behaviors, may purport to serve the dramatic function of communicating that when Trip asked Grace to marry

him she wasn't really ready, while simultaneously communicating that they are both getting more upset and that Grace currently feels disaffiliated with the player. The associated beat code may simultaneously describe the author's vision of the desired run-time experience, and, when executed, implement the author's vision of the desired runtime experience. But when the author, or another player, plays the experience, Trip and Grace actually seem less upset than in the preceding beat, even though they are supposed to be more upset. What happened here is that the details of the writing, and how the details of their physical performance actually read, are *extra-architectural*; they lie outside the literal code of the system. Even though the beat is "performing to spec", other sign systems are subverting its interpretation. Every AI system is doubled. A description of the code system is not enough – we need to examine the rhetorical system.

4.2 The Rhetorical System

The signs of both $system_1$ and $system_2$ are multi-articulated; their meaning arises both from syntagmatic and paradigmatic constraints established by the respective code systems, but also from a collection of sign systems *outside* of the code systems. This collection of external code systems is the rhetorical system. Both authors and audiences make use of the rhetorical system in narrating the operation of the system and forming intentions with respect to the system. The code and rhetorical systems are tightly entangled; both play a role in understanding interpretive and authorial affordances.

4.2.1 (Audience) Interpretive Surplus

$Syntagm_1$ never completely describes all the properties of $syntagm_2$; though $system_2$ literally prescribes the possible elements (paradigm) and spatial and temporal relationships between elements (syntagm) of $syntagm_2$, a portion (perhaps a large portion) of the signification is determined by external sign systems. This interpretive surplus occurs because $system_2$ operationalizes a meta-language ($syntagm_1$) for describing the audience experience ($syntagm_2$). The signifieds of this meta-language are themselves signs, participating in external sign systems, which are handled by the meta-language.

The crafting of these external, handled signs, becomes an irreducible problem in design and aesthetics. These handled signs must be crafted to marshal the signifying resources of these external sign systems in such a way as to match the purported meanings of the code system. For example, in *Façade*, we as authors have to write dialog that consistently communicates the character of Grace and Trip, while communicating meanings appropriate for a specific beat goal within a specific beat, while also being re-sequenceable to various degrees. Specific lines of dialog must meet multiple constraints established by how the code machine will make use of the line. Additional meaning is carried by how a voice actor performs the line. The nuances of emotional tone, irony, sarcasm, desperation, etc., communicated by the voice performance, must also be consistent with these constraints. In authoring *Façade*, there is a reciprocal process between authoring these handled signs (e.g. dialog, snippets of animation data) and code-level authoring within the architecture. Consistency between handled signs and manipulation by the code machine is established by moving back and forth in the authoring of these two domains. But consistency is not the same as identity; there are always aspects of audience interpretation that escape the code machine.

Another avenue for interpretive surplus is connotation; the handled signs may become the plane of denotation for a connotative system. For example, in *Terminal Time*, the ideological arguments made by the system are often (purposely) undermined through irony. The details of imagery, music, and the narrative track connote irony, while at the level of denotation an earnest argument is being made. For example, if the anti-religious rationalist ideologue has been activated, a 20^{th} century event it may make use of is the Chinese invasion of Tibet. Within the knowledge base, the two actors of this event are Tibetan Buddhists (which the system infers are a kind of Religious Group), and Maoists (which the system infers are a kind of Rationalist through their connection to Marxism). Furthermore, the event is a War, instigated by the Maoists (Rationalists) against the Buddhists (Religious Group), in which the Maoists are successful. This is enough for the Anti-Religious Rationalist to decide it can use this event as a Positive Example of Rationalist Progress. Assuming that this event spin (the ideologically-slanted representation of the "objective" representation in the knowledge base) makes it into the final generated documentary, the system will earnestly argue that this is a positive example of Rationalists mopping up the remaining dregs of irrational religion (e.g. "There were reports that Buddhists monks and nuns were tortured, maimed and executed. Unfortunately such actions can be necessary when battling the forces of religious intolerance.") over a montage of Tibetan Buddhist imagery and Chinese soldiers holding monks at gunpoint, while playing the happy, "optimistic" music loop. The system does not "know" that it is undermining its argument through irony; irony is not a property described within the code machine. We as system authors marshaled the handled signs (language, video clips, music) to connote irony on top of the structure explicitly provided by the code machine.

Given that the audience interpretation of $syntagm_2$ always escapes full specification by the code machine, it may be tempting to conclude that computer-based art practice should primarily make use of the signifying resources of external sign systems via handled signs. Crafting the handled signs, animation snippets, imagery, video clips, music loops, and so forth, falls comfortably in the realm of more traditional art practice. Such an approach would move back towards the "code as a hack" model, throwing together the minimum code machine necessary to coarsely manipulate handled signs. But this approach would severely compromise the intentional affordances. As the interpretive surplus becomes larger and larger, with more of the interpretive affordance pushed onto the handled signs, an imbalance grows between the intentional affordances offered by the system and the system's ability to actually respond to these intentions. The rich handled signs suggest many avenues of action to the audience. But with no corresponding richness in the code machine, there is no way for the work to respond to these actions; the rich, coarsely handled signs suggest a richness of response that the work can't satisfy. But the reason for designing a rich and expressive architecture goes beyond the "utilitarian" goal of supporting audience agency. The architecture ($system_1$), and systems designed within it ($syntagm_1$), are themselves embedded in a meshwork of external sign systems, providing the AI-based artist with a rich architectural surplus.

4.2.2 Architectural Surplus

Agre [2] describes how AI technical practice provides narrative affordances that support AI researchers in creating stories describing the system's operation.

> … the practical reality with which AI people struggle in their work is not just "the world", considered as something objective and external to the research. It is much more complicated than this, a hybrid of physical reality and discursive construction. … Technical tradition consists largely of intuitions, slogans, and lore about these hybrids, which AI people call "techniques", "methods", and "approaches"; and technical progress consists largely in the growth and transformation of this body of esoteric tradition. [2:p. 15]

Different practices (e.g. classical AI, interactionist AI) provide different affordances for narrating system behavior. For the classical AI researcher, the discursive construction consists of ways of talking about "goals", "plans", and "knowledge", while for the interactionist AI researcher, the discursive construction consists of ways of talking about "embodiment", "action", and "improvisation". These discursive constructions are a necessary part of the functioning of the system.

> To understand what is implied in a claim that a given computer model "works", one must distinguish between two senses of "working". The first, narrow sense, again is "conforms to spec" – that is, it works if its behavior conforms to a pregiven formal-mathematical specification. … the second, broad sense of "working" … depends on specific words of natural language. As I mentioned at the very beginning, an AI system is only truly regarded as "working" when its operation can be narrated in intentional vocabulary, using words whose meanings go beyond mathematical structures. When an AI system "works" in this broader sense, it is clearly a discursive construction, not just a mathematical fact, and the discursive construction succeeds only if the community assents. [2:p. 14]

In typical AI research practice, these affordances are often not consciously acknowledged or manipulated. Rather, they serve as part of the unconscious background, co-evolving with the technical practice as a silent but necessary partner in the research. Systems are spoken of as having "goals" or engaging in "embodied action", as if these were primitive, readily detectable properties, like being blue, or being cold, rather than the hard-won results of rhetorical construction and debate. But in Expressive AI practice, these discursive constructions are an explicitly manipulated resource, an architectural surplus that makes the architecture not just a bunch of code, but a way of thinking about the world.

Within the semiotic framework of this chapter, the architectural surplus (an interpretive surplus on the author side), can be understood as one or more meta-languages, in which the signs in $system_1$ ($syntagm_1$) form the content plane, and as one or more connotative systems, in which signs in the meta-language form the plane of denotation.

For example, consider joint goals in ABL. The code sign for a joint goal appears in Figure 5. The sign signifies that a team of ABL agents will attempt to achieve `Goal1()`. A meta-language allows us to talk about and thus operate on these code signs. This meta-language consists of ordinary language that has been co-opted into talking about code signs. This meta-language in turn serves as the plane of denotation for a connotative sign system – this connotative sign system contains the "spillover" of the co-opted ordinary language, connotative meanings that escape the strict meaning of the code signs. In this case, the meta-language sign for a joint goal connotes the idea of a team of people working together, with all the non-formalized richness of this notion. The connotation lifts the code sign out of the circumscribed meaning provided by the architecture, and into the more open-ended sign system used to talk about coordinated human activity in the everyday world. Once lifted into this connotative system, the author can use the connotative sign system to think about the human realm of teamwork. But new signs reached by thinking in the connotative plane can in turn have signifiers in the meta-language whose signifieds lie back in the code system. Thus ordinary language, in this case the ordinary language of human teamwork, becomes a meta-language for talking about and manipulating a technical system, in this case the code system for joint goals in ABL. This movement, from code system, into ordinary language, and back into code system, creates a circulation of signs that suggests both new ways of using the architecture and new architectural elaborations, in this case news ways of using joint goals and new architectural elaborations for joint goals.

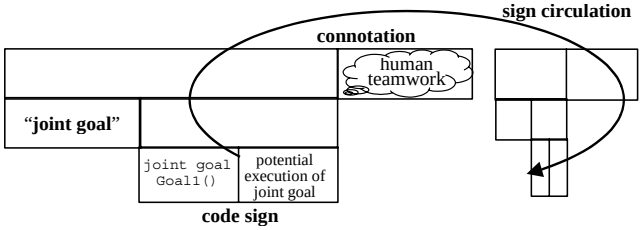


Figure 5. Code signs, meta-language, and connotation.

Consider first how the ordinary language system of human teamwork suggests new ways of using joint goals. In the everyday human world, we think of people coordinating to achieve goals they want to achieve; that is, we imagine people having a positive emotional valence towards a goal. Two people might team up to hang a picture, or change a tire, but we don't picture people teaming up to have a big fight, or teaming up to accomplish a mugging, with one team member the victim and one team member the mugger. An author may thus never think of using joint goals to coordinate a big fight among two agents. But now imagine that in the connotative plane we start thinking about teams of movie actors or stage actors. In acting within a play or movie, human actors often tightly coordinate in the carrying out of dramatic activity in which the characters strongly oppose each other, as in, for example, a play in which a marriage falls apart as years of buried frustrations and misunderstandings are revealed. Now this gives us the leverage (meta-language) to imagine using joint goals in *Façade* to tightly coordinate conflicts between characters. Ordinary language, used as both the plane of connotation and as meta-language, is a necessary part of the *total* system – it provides the code system with broader meaning and consequently suggests new ways of manipulating the code system. Note that this example involves consciously manipulating and exploring the plane of connotation in order to reveal a new

possibility within the code system. If we were uncritically wedded to the ordinary language system of "rationality", in which people only pursue goals for things they emotionally desire, then the code system idea of jointly accomplishing conflict may never arise.

The plane of connotation and meta-language not only suggests ways of using the code system ($syntagm_1$), but modifications and elaborations of the architecture itself ($system_1$). Continuing with the joint goal example, consider the control of activity within a joint goal. In ordinary language, when we imagine team members accomplishing a task together, we often imagine the decision of what step to do next being distributed among the team members. Certainly there are hierarchical situations in which a team leader is responsible for managing the teams, but many teamwork situations are more collaborative and decentralized. Now consider the initiation of joint goals in the code system. When one member of a team initiates the joint goal, the other members of the team, on successful entry into the joint goal, spawn the goal at the root of their active behavior tree (ABT). Only the joint goal initiator has the goal deeper within the ABT. If other members of the team initiate joint subgoals in the service of the original joint goal, these joint subgoals will appear at the original initiator's ABT root. This is a bit counter-intuitive, given that within the ABT subgoals are normally children of the goal (via a behavior) they are in service to. But strictly at the code level there is nothing wrong with this arrangement. However, consider how the ABT is connotatively read or interpreted. The ABT captures the structure of an agent's thoughts, its mind. It is not just a bookkeeping mechanism controlling execution, but a *representation* of the agent's activity. Reflective processes (meta-behaviors) may treat the ABT directly as a representation. But even without reflection, the mechanisms for success and failure propagation, the many annotations that modify success and failure propagation, and continuously monitored conditions, all work together to support the reading of the ABT as a representation. When a goal appears deep in the ABT, it is enmeshed in more complex patterns of activity than a goal shallower in the ABT – ABT depth becomes a proxy measure for the complexity of the agent. With this reading of the ABT, combined with the ordinary language model of teamwork, the default joint goal initiation mechanism is seen as lacking. Initiated joint goals, since they are always at the root of the ABT, aren't able to fully participate in complex patterns of activity. This is particularly problematic for "flat" teams, in which all team members equally participate in the control logic for the team, and thus both initiate and respond to requests to enter joint goals. This circulation between readings of the ABT, code signs for joint goals, and readings of these code signs, suggests an architectural modification supporting the initiation of joint goals anywhere in the ABT.

Authorial affordance consists not just of the code system relationship that $syntagm_1$ simultaneously implements $system_2$ and describes $syntagm_2$, but also of the rhetorical relationship that $syntagm_1$ is readable and handleable by interpretive systems and meta-languages. An architecture is a machine to think with. The complex circulation between code signs and the interpretive framework provides authors with both resistance (some things will appear hard or impossible) and opportunity (new ideas arise). Thinking with the architecture suggests new audience experiences, creating a feedback loop between authorial intention and the details of the *total* system (code + rhetoric). But establishing this interpretive framework, the plane of connotation and meta-language, takes real work. It is the outcome of a practice that simultaneously tries to articulate the code machine *and* the ways of reading it and talking about it. In contrast, a practice that views the system as a hack, as a means to an end, will likely construct systems with poor authorial affordances, lacking both the code system relationships and rich rhetorical frameworks necessary to enable new audience experiences.

4.3 Idioms

Idioms are ways of using an architecture, conventional structures for the authoring of $syntagm_1$. Idioms arise through the interplay of the architecture and its interpretive frameworks. In a sense, the idioms actually cash out the interpretive framework, being the place where interpretation and code meet. This is why idioms are so important for truly understanding an architectural system. An abstract description of a code system will make use of all kinds of ordinary language words, such as "plan", or "embodied activity", or "learning", but understanding the particular entanglement of rhetoric and code that is the total system requires examining the detailed circulation between these language signs and code signs. Idioms are the place where this detailed circulation occurs.

As idioms become larger and more diffuse, they begin restricting the circulation between code and rhetoric. The code signs become large and diffuse, making the connotative lifting and meta-language handling difficult. Idioms can thus reveal breakdowns in the total system, conceptual domains in which the circulation between rhetoric and code are restricted. The breakdowns suggest architectural opportunities, modifications of the architecture that enable new idioms and simultaneously re-articulate the interpretive sign systems, providing new ways of talking and thinking about the code system. Systems built without an explicit concern for authorial affordances are likely to be *all* idiom, and thus severely restrict the circulation between rhetoric and code. This would be the case, for example, if *Façade* was written as a giant program in a standard programming language such as C. The only code signs at our disposal would be the rather low-level signs provided by C. Everything else would be idiom, with large chunks of C code having only a diffuse relationship to signs of the audience experience ($syntagm_2$) and to connotative and meta-languages. This extreme case of the code system being nothing but idiom, code piled on code, provides poor authorial affordances, making it difficult to think about, discover, and express, new conceptual frameworks and new audience experiences.

4.4 Generality of the Doubled Machine

The use of a structural semiotic terminology in this chapter, with the focus on "sign systems", "languages", "connotation" and so forth, may lead a reader to conclude that the analysis of affordances in terms of doubled machines of rhetoric and code is only useful for classical AI systems, with their explicit focus on symbolic knowledge. The analysis applies much more broadly than this, however, to any AI or ALife practice. All such practices make use of a rich entanglement between technical systems and ways of talking and thinking about the technical system. Consider a robot built along the lines of subsumption architecture [6], in which finite state machines mediate rather directly between sensory input and motor actuation. The finite state machines may in fact be implemented entirely in hardware, rather than as code in a general purpose micro-controller. Yet there is a still a "code

machine" that participates in complex discursive constructions. Wires bearing voltages are in no less need of interpretation than fragments of textual code, and participate in the same sign system relationships that support interpretive and authorial affordances.

The focus in this chapter on authorship may similarly lead a reader to conclude that this analysis is not relevant to machine learning. But again, the methods of machine learning consist of a technical/rhetorical system, one organized around the "learning" or "discovering" of "patterns" in "raw data". But, of course, human authors select the primitive features, define the representations of hypotheses or distributions, define the search methods employed to tune parameters, and design how particular machine learning methods are embedded in larger architectures. For example, *Office Plant #1* makes use of the technical/rhetorical system of text learning as part of an architecture supporting the creation of a non-human companion responding to email activity.

5. Conclusion

This paper develops authorial and interpretive affordances as central terms in the hybrid practice of Expressive AI. The relationship between these two affordances shows how Expressive AI is simultaneously concerned with art's creation of meaningful experience (and the consequent focus on interpretation of the art object), and AI's construction of machines that can be understood as behaving intelligently (and the consequent focus on the structures, properties and processes of these machines). Structuralist semiotics, through its concern with sign systems and the relationships between systems, provides a common ground in which both the artwork as experienced by the audience and the construction of machines as experienced by the author can be seen as instances of sign systems – this provides the framework for a more detailed analysis of the relationship between these affordances.

As an analytical framework, structuralist semiotics has its limits. Arising from the tradition of Sassure, its view of the world as a meshwork of language systems whose rules can be analyzed has trouble accounting for the actual processes involved in the use and production of signs. Some work in the analysis of computational media has fruitfully made use of Peirceian semiotics, whose sign concept includes a notion of meaning more amenable to process (e.g.[1, 7:chapter 4]). Further analysis of the negotiation of meaning in technical systems could fruitfully make use of ethnographic and phenomenological frameworks. However, the structuralist analysis here, with its focus on the relationships between sign systems, goes a long way towards understanding both how and why Expressive AI is simultaneously concerned with the code system and audience interpretation.

6. REFERENCES

[1] Anderson, P. B., Holmqvist, B., Jensen, J. F. 1993. The Computer as Medium. Cambridge: The Cambridge University Press.

[2] Agre, P. 1997. Computation and Human Experience. Cambridge, UK: Cambridge University Press.

[3] Barthes, R. 1972. Mythologies. (Trans: Annette Lavers). New York: Hill & Wang. Translation of Mythologies, first published in 1957.

[4] Barthes, R. 1967. Elements of Semiology. (Trans: Annette Lavers & Colin Smith). New York: Hill & Wang. Translation of Eléments de Sémiologie, first published 1964.

[5] Boehlen, M., and Mateas, M. 1998. Office Plant #1: Intimate space and contemplative entertainment. Leonardo, Volume 31 Number 5: 345-348.

[6] Brooks, R. 1986. A robust layered control system for a mobile robot. IEEE Journal of Robotics and Automation, RA-2(1).

[7] Frasca, G. 2001. Videogames of the Oppressed: Videogames as a Means for Critical Thinking and Debate. Masters Thesis, Interactive Design and Technology Program, Georgia Institute of Technology. Available at: www.ludology.org.

[8] Gibson, J. 1979. The ecological approach to human perception. Boston: Houghton Mifflin.

[9] Hjelmslev, L. 1961. Prolegomena to a Theory of Language (trans. Francis J Whitfield). Madison: University of Wisconsin Press. Translation of Omkring Sprogteoriens Grundlæggelse, first published in 1943.

[10] Mateas, M. 2002. Interactive Drama, Art and Artificial Intelligence. Ph.D. Thesis, Computer Science Department, Carnegie Mellon University. CMU-CS-02-206.

[11] Mateas, M. 2001. Expressive AI. Leonardo: Journal of the International Society for Arts, Sciences, and Technology, 34 (2), 147-153.

[12] Mateas, M. and Stern, A. 2003. Integrating Plot, Character and Natural Language Processing in the Interactive Drama Façade. Proceedings of *Technology for Interactive Digital Storytelling and Entertainment* (TIDSE) 2003.

[13] Mateas, M. and Stern, A. 2000. Towards Integrating Plot and Character for Interactive Drama. In *Working notes of the Social Intelligent Agents: The Human in the Loop Symposium.* AAAI Fall Symposium Series. Menlo Park, CA.: AAAI Press.

[14] Mateas, M., Vanouse, P., and Domike S. 2000. Generation of Ideologically-Biased Historical Documentaries. In Proceedings of AAAI 2000. Austin, TX, pp. 236-242.

[15] Newell, A. 1982. The Knowledge Level. Artificial Intelligence 18: 87-127.

[16] Newell, A. and Simon, H. 1976. Computer science as empirical enquiry: symbols and search. Communications of the ACM, 19:113--126.

[17] Norman, D. 1988. The Design of Everyday Things. New York: Doubleday.

[18] Penny, S. 2000. Agents as Artworks and Agent Design as Artistic Practice. In K. Dautenhahn (Ed.), Human Cognition and Social Agent Technology. Amsterdam: John Benjamins.

[19] Saussure, F. 1974. Course in General Linguistics. London: Fontana. Translation of Cours de linguistique generale, first published in 1916.

[20] Turing, A. 1950a. Computing machinery and intelligence. Mind 59:433-60.

Contact Expressions For Touching Technologies

Kevin McGee
Department of Computer and Information Science
Linköping University
58 183 Linköping
Sweden
kevmc@ida.liu.se

Annika Harup
IT University of Göteborg
Box 8717
40 275 Göteborg
Sweden
It2haan@ituniv.se

ABSTRACT
Human-computer interfaces typically involve limited tactile input and audio/visual output, and even when the interface has been enhanced with speech, gesture, and haptics, this is often done to supplement (or compensate for) audio/visual output. Even the most interesting and elaborate work on force-feedback does not seem to be yet envisioning the full expressive potential of physical contact, in particular, the "contact expressions" used routinely by people and animals in different contexts. People use contact expressions when other forms of communication are inappropriate or impossible, to supplement other forms of communication, or because the physical contact itself has significance. As robotic toys and embodied technological devices become smaller, more portable, more durable, and more commonplace, it is our belief that contact expressions will become an important area of interface design and will open up new areas of study for applied semiotics. This paper describes a "contact cushion" we have built and used to explore some of the potential for contact expressive devices – and outlines a preliminary design taxonomy of basic contact expressions.

Keywords
Contact Expressions, Semiotics, Tactile interfaces, Haptics

1. INTRODUCTION
SCENARIO: *Karo moves into the living room to find Janet lounging on the sofa with her infant son, Paul, who is sleeping beside her. Young Paul starts to thrash in his sleep, so Karo crawls up and sprawls next to him. Janet pats Karo affectionately on the head and Paul opens his eyes, so Janet then pats him on the head, too. Paul smiles sleepily at the joke, wraps one arm around Karo, and snuggles into the warmth; the two of them start to breathe deeply and evenly in tandem contact. A little later Karo indicates that Janet's friend John is calling, but Janet can tell from the signal that he is in a chatty mood and so decides not to answer.*

Karo puts his nose under Janet's arm, nudging it up a few times until she gets the idea that he wants to be stroked; he settles his head into her lap. When Karo's head begins to get too heavy, Janet tries to push it off; but Karo notices that at that moment he would disturb the sleeping baby if he moved in the direction she is pushing, so he resists – and when Janet stops pushing, he moves his head in a different direction. After a while, Janet drifts off to sleep – and wakes when the nearly-empty bowl of chips that was beside her slides off the sofa and spills onto the carpet. She reaches out her hand to confirm that Karo is still lying beside her, then pats/pushes him to indicate that he should take care of it; he responds by getting up, licking/eating up the spilled chips, and picking up the bowl in his mouth and taking it into the kitchen.

In this not-too-distant-future scenario, Karo is a robot – and all of this takes place silently in the dark.

The purpose of this scenario is to highlight some of the ways that signifying physical contact is an almost unnoticed part of everyday life – and to suggest ways in which such "contact expressions" may be fruitfully incorporated into future computational devices.

What we mean by "contact expressions" is probably most evident in the range of examples from the scenario, but we can also say that by our view they involve mutual signification, are not simply "parasitic" on natural language (nor entirely subsumed within it), and include both instrumental and experiential touch. In other words, just as "facial expressions" involve significance beyond the control and display of certain muscles, so, too, "contact expressions" involve significance beyond simple properties of touch. And just as there is an emerging vocabulary and set of techniques for designing graphical interfaces (GUIs), we wish to contribute to the emergence of a similar vocabulary and set of techniques for designers of contact expressive interfaces (CUIs). In our view, contact-expressive design can involve anything from simple "massage chairs" to more complex, pro-active feedback devices suggested by the description of Karo. To paraphrase John Austin, we would like to begin addressing the problem of "How to (design computational devices that) Do Things With Contact Expressions."

2. SURVEY
In general, there seems to be very little direct research and development on contact expressive devices. What follows is a short survey of work in three relevant traditions: studies in the behavioral sciences that attempt to show the significance of "touch"; attempts to create taxonomies (or "vocabularies") of physical touch; and the development of technologies that use physical contact as a significant part of the interface.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

2.1 Behavioral Sciences

Although there is a fairly large body of research into what may be called "nonverbal communication", the majority of this work is on visual expression (facial expressions, physical appearance, direction of gaze, physical posture, orientation, movement, proximity and other visible forms of "body language") and paralinguistics (i.e., vocal inflection, pitch, volume, speech rate, etc.). A representative definition of what is meant by non-verbal communication makes this clear: "Nonverbal communication, then, could be defined as that part of a message, which is not words, but which may accompany words or occur separately from words – and includes facial expressions, gestures, posture, spacing, tone of voice, pitch, volume, speed of talking, etc." [22]. Although a typical survey of nonverbal communication will include references to studies of touch, there is usually very little detail; and a survey of the literature on nonverbal communication also reveals that only a small percentage of it is devoted directly to studies of touch.

There does seem to be widespread acknowledgement of this neglect by the researchers who study nonverbal communication, and one suggested explanation [20] is that although this may partly be the result of touch-related inhibitions and taboos, it is almost certainly also related to the methodological difficulties of studying physical contact. These difficulties include the fact that casual contact among strangers is rare; intimate contact between friends or lovers is usually private; and the factors relevant to touch are difficult to isolate from other senses and forms of expression without creating further research challenges. All of these difficulties also help to explain why most of the research on touch is in the form of observational correlation studies. (For a survey of this research, see [19].)

Finally, within the field of psychology there is the concept of "contact gestures" which are "physical commands" used instead of visible gestures. So, for example, a young child who wants an adult to open a container may use a contact gesture to actually place the adult's hands on the container and put them through the motions of opening it. There has been some work comparing the use of contact gestures among autistic and non-autistic children and apes [13], but in general, there does not seem to be an existing descriptive taxonomy of contact gestures.

2.2 The Semiotics of Touch

Although semiotics in the large is concerned with the full range of "signification" (or "meaning"), physical contact is one of the areas that seems least explored by semioticians. To be sure, it is indicated (usually in a cursory manner) as part of the larger field of semiotics, but there seems to be very little detailed study, theory, or discussion about the signifying nature of physical contact.

Introductory texts on semiotics usually devote no more than a line or two to the effect that physical contact involves "body codes" (body contact, proximity, physical orientation, appearance, facial expressions, gaze, head nods, gestures and posture) – and "group-specific codes" (gender, age, race, culture, status, etc.). As indicated above, many of these codes are what behavioral scientists typically consider part of nonverbal communication (rather than physical contact *per se*). Even the most encompassing scholarly discussions or surveys of semiotics [16] do not seem to have included much detailed investigation or analysis of physical contact.

And unlike the behavioral sciences, this does not seem to occasion as much comment among practitioners. We can imagine at least two possible contributing factors: the first is that semioticians tend to come from linguistic or philosophical traditions (rather than, say, physical therapy); the second is that many semioticians may still, in many ways, believe in the validity of abstract taxonomies – in which case the "physical embodiment" of signifying systems may seem irrelevant to the signification processes themselves (and are therefore subsumed under more general semiotic taxonomic categories). There is indeed suggestive evidence for this second hypothesis if we look at the kinds of semiotic descriptions that have been proposed for physical contact. So, for example, there have been some proposals for large taxonomies of types of bodily contact; classifications of bodily contact common in Western Culture; a five-level taxonomy of type/meaning pairings; and proposals for modeling the possible relationships between physical contact and status (see [19] for an overview).

On a related note, although Piaget and his colleagues have done a great deal of research studying "sensori-motor intelligence" – and Piaget himself was deeply interested in integrating semiotics into his model of cognition [24] – the literature on development psychology does not appear to contain studies of what might be called the "physical semiotic function." A contributing factor may be the structure of Piaget's stage model, which proposes that the semiotic (or "symbolic") function arises as part of the transition out of the sensori-motor stage. In other words, the semiotic function is seen largely in terms of the emergence of (referential) language and *thought* – by contrast to the more "primitive" sensori-motor reflexes. By this view, the emergence of language is related to the observation that words (and "thoughts") can "represent" physical actions that need not be taken. The fact that there can be physical "puns" – that some contact expressions can stand for, or reference, *other contact expressions* – does not seem to have been the subject of study.

2.3 Tactile Technologies

Although some aspects of haptic research are as old as any work on human-computer interaction, in many ways touch is still a nascent field of HCI study. There are, of course, a number of efforts to develop technologies that accept tactile (or haptic) input and, to a lesser degree, provide tactile output ("force feedback", etc.). Most of the work to develop "tactile interfaces" has concentrated on simulating different aspects of the physical world (game force-feedback devices, motion simulators, and the like); on providing additional feedback in multimodal systems for complex tasks (force-feedback for molecular docking systems, etc.); or on "mapping" the movement of an on-screen pen over physical contours of a surface to a (force) feedback input device. (For a summary of the history of research on haptic and force-feedback interfaces, see [28].)

One of the earliest projects to use tactile conviviality as part of the interface was in the creation of Noobie, The Furry Computer [10]. It was a computer in the form of an over-sized, plush toy for small children to climb up on, snuggle into, sit (or lay) on, etc. This theme, of using plush toys to make the interface friendly

and familiar, has been elaborated by a number of subsequent researchers [11,12,15,17]; however, very few projects seem to be elaborating the physical contact aspects of the original Noobie project.

A sampling of current projects gives an indication of both utilitarian and experiential applications of touch-oriented computational research. There are projects to develop telephones with various haptic qualities [4,21]; to develop various kinds of "musical objects" that respond to touch [27,31]; to develop physically embodied interface objects ("tangible bits") that can be manipulated in various ways [14]; to develop contact-responsive plush electronic toys (Furby, Barney, Tickle Me Elmo) [7,18,29]; and to develop hand-held devices with "ambient touch" interfaces [25] that can support different physical interactions such as tilting, dropping, spinning, rubbing the device against something else, and so on. There is also related work in the field of textile design, where online shopping is motivating research into methods for simulating tactile aspects of different materials [8]; and in recent years, there have been semi-serious proposals for combining VR and force-feedback for "virtual sex" [3].

3. PROBLEM STATEMENT

As indicated in the survey, there does not seem to be much in the way of either semiotic analysis of – or applied research into – the extended possibilities of what we are here calling contact expressive devices. In some cases the absence of research is simply curious. For example, most human-computer interaction already involves what could be considered contact expressive input: we type at keyboards, move mice, draw with pens on tablets, and use pressure-sensitive screens. And although this is obviously a very limited subset of the full potential of contact expressive interfaces, it is almost inevitable that haptic interaction will be extended in various ways (not the least of which will be to make more expressive use of haptic *output*.)

In other cases, the absence of an explicit emphasis on contact expressions is actually startling. There are, for example, many robotics projects around the world with the stated intention of improving life for the elderly, the autistic, and the allergic; however, it is rare for these projects to include physical contact with the robots as an explicit goal (for some important exceptions, see [5,9]). Even project reports devoted to robotics and "social interaction" tend to ignore almost all aspects of physical contact. So, for example, in the otherwise excellent survey article on the many dimensions of "embodied social interaction" [11], there is literally no mention of physical contact.

Beyond these limitations, there is also the almost total absence of any semiotic study of contact expressions – and virtually no work of which we are aware that attempts to leverage such analyses in the service of designing and implementing useful and experientially interesting contact expressive devices.

We feel, therefore, that there is still major work to be done studying, understanding, and making use of contact expressions. Our particular focus in this paper is to contribute to the development of this area of study by sketching an initial design taxonomy and an example contact-expressive design Pattern.

4. METHOD

Since the application of semiotics to computational interfaces is still in its infancy, a brief description of our orientation may help the reader contextualize the work we report here. Our work can be understood as part of the larger effort on the "syntonic design" [23] of empowering computational devices, applications, and services. Briefly, this consists of a cycle of implementation, study, and revision. The implementations are based on cognitive insights from the intersection of developmental ("constructivist") psychology and semiotics [24,2]; the studies are largely qualitative in which insights about *the activities of participants using particular implementations* form the basis for cognitive hypotheses, the development of design vocabularies and guidelines [1], iterative modifications of the implementations, and further study.

In order to begin elaborating a vocabulary for contact expressions for computational devices, two small, exploratory studies with a contact expressive cushion (the PillO'Mate) were conducted. The first study explored the use of the cushion as a therapeutic presence, and illustrates a number of issues relevant to understanding the *experiential* (or "self-fulfilling") potential of contact expressive devices. The second study explored people's reactions to using the cushion as a telephone "pager", and illustrates some issues relevant to their *instrumental* potential.

These studies involved qualitative explorations of contact expressions and some of their significance ("what do different kinds of contact expressions mean to you?" and "why do you associate a specific contact expression with a specific type of person?", "if we changed the device so that it was contact expressive in this other way, what kind of person would you assume is trying to contact you?" etc.). These discussions are the basis for formulating some initial descriptions of human contact acts – and for then incorporating them into an evolving "design Pattern Language" for contact expressive devices.

In the work reported here our goal was to complete an initial cycle of implementation, study, and analysis. In other words, we felt that there was very little known about contact expressions and which aspects might be relevant to the design of contact expressive devices. Therefore, in order to get some idea of which areas are promising for further, sustained investigation, we chose to focus on a quick implementation and a series of short exploratory studies to get some initial indications of how people used them and what significance was ascribed to them. Although it will be important later to do more rigorous studies, it simply felt premature to develop testable hypotheses, do extensive studies with control groups, or gather quantitative results. The cushion, the studies, Speech Act Theory, and Pattern Languages are now described in more detail.

4.1 PillO'Mate Device

The PillO'Mate is an oversized cushion, inside of which is the following technology: a fine wire-mesh proximity-sensor, a GSR (Galvanic Skin Resistance) sensor, a vibration motor, a heating pad, and a speaker. The electronics inside the cushion are surrounded with "granular" Tempur™, which was chosen so that the cushion would to be not too heavy, it would be pleasant to hold and hug, and yet it would still protect the electronics inside. The back of the cushion is covered in fleece (to conduct heat),

and the front is covered in imitation leopard-skin to encourage stroking, cuddling, and nuzzling. (For the studies reported here, it was not possible to completely integrate all the necessary components directly into the cushion. The resulting solution has a "tail" which consists of cords that connect the heating pad to an external transformer for the heating pad, an amplifier for the speaker, and a PC for the audio clips. Although this tethered solution made certain kinds of studies impractical, it also made some users explicitly aware of and interested in the potential of more mobile contact expressive devices.)

People use the PillO'Mate in much the same way they do an ordinary large cushion – they pick it up, put it on their laps (or hug it to their chests), fiddle with it, and lean, lay, or put their heads on it. The cushion switches to an active state when it detects an approaching object (such as a hand) within 5cm, but only by being stroked or touched in specific ways will it react with sound, vibrations, and/or changes of temperature. (For more detail, see [6].)

4.2 PillO'Mate Studies

Study 1: Warming Up to Contact Expressions. The original impetus behind the PillO'Mate was to develop a device that helped people relax. Using a pet (cat) as an initial source of inspiration, the device was developed as a way to explore certain qualities that people find relaxing about having a cat sleeping on their laps. It is important to emphasize that the goal was not to simulate a real cat, but rather to explore and support different aspects of touching, feeling and listening. The first study, then, is based on observations of people (many of them attendees at Comdex 2003 in Gothenburg) sitting with the PillO'Mate and talking about it in terms of its soothing qualities.

Study 2: Contact Expressive Pager. In order to explore some of the instrumental potential of contact expressions, we did a very small study of people's concerns by discussing scenarios involving the PillO'Mate as a "contact expressive pager." This consisted of some exploratory meetings with six people in which they were asked to sit with the cushion on their laps, close their eyes, and imagine that when the cushion started vibrating it was an indication that someone was calling. They were then asked to describe characteristics of the imagined caller based on the characteristics of the contact expression. The study was divided into two parts: in the first part, participants discussed the significance of the current PillO'Mate for different categories of people calling, types of call, etc.; in the second part, they discussed how they felt the contact expressions of the PillO'Mate could or should be changed in order to better match the their experience and expectations of different callers, phone calls, and so on.

4.3 Applied Semiotic Analysis of Contact Expressions

There are a number of possible semiotic models that could be used as the basis for enriching our understanding and application of contact expressions. Since we are interested in exploring the co-adaptation and co-evolution of cognition and mediating technology, we pursued an analysis of *activity* inspired by the earlier phenomenological work of Winograd and Flores [32]. That is, we are trying to understand *contact acts* – the first-person "how and why" of contact expression signification.

As a further elaboration of Speech Acts, we are also in the initial stages of trying to build a design vocabulary for contact expressive devices.

4.3.1 Speech Act Theory

In order to frame the results and insights of our work, we now provide a brief overview of (a slightly generalized version of) Speech Act Theory (for a more detailed overview, see [32,26,30]). Speech Act theory emphasizes what people want to *do* with language (as opposed to trying to establish the truth-value of what any particular statement signifies), and proposes three major types of "acts": locutory, illocutory, and perlocutory. (In order to provide relevant examples below, we reference "contact acts" which we will more fully elaborate later in the paper.)

The locutory dimension of a communication act has to do with its material generation – locutory differences in a contact act includes whether the contact expression is soft or hard, whether it vibrates or not, the rate of vibration, and so on. The illocutory dimension of a communication act has to do with the *intention* of the act – illocutory differences in a contact act include whether the physical contact *asserts* ("someone is calling"), whether it *expresses* ("I am happy"), whether it *directs* ("move over there"), whether it *interrogates* ("what is this thing?"), and so on. The perlocutory dimension of a communication act has to do with the effects it can have on the receiver's actions, beliefs, or judgments – perlocutory differences in a contact act include whether *the person being touched* calms down, has a change of mind, is physically displaced by the contact, leans into it, and so on.

In addition, an important aspect of the Speech Act model is the emphasis on the *felicity* conditions of an act – that is, on the conditions that make such acts appropriate. For this study, we consider this to be largely synonymous with Alexander's notion of "context" (see below).

Some limitations of this model will become clear as we start to apply it to the analysis of instrumental and experiential contact expressions. Here we briefly highlight two of them: first, it is not clear whether all illocutory dimensions of speech acts are present in contact expressions (nor whether the illocutory categories of speech acts account for all aspects of contact expressions); and second, it is not clear how the current instrumental emphasis of the model can account for such acts as jokes, play, altruism, creative expression (singing, poetry), and other self-fulfilling forms of expression.

4.3.2 Pattern Languages

Christopher Alexander developed Pattern Languages as a way to effectively identify and describe "good" architectural solutions – solutions that satisfy both structural and experiential requirements.

For our purposes, Patterns are important because they describe an effective (architectural) *feature* that resolves certain conflicting *forces* in some *context* – and they do so in a way that helps a designer "know what to do." So, a classic example is Alexander's proposal for (natural) "Light on Two Sides" as a pattern that needs to be addressed in order for a room to be "good." This example illustrates that Patterns are specific, flexible, sharable, testable, and pragmatic. This particular

Pattern is specific enough for anyone to be able to determine whether or not it is satisfied for a particular room; it is flexible enough that it may be satisfied many different ways; it is sharable in that it can clearly be described for someone else; it is testable in that people can determine whether or not the proposed feature does, in fact, improve the quality of rooms where it is present; and it is pragmatic because the description of the pattern helps a design know what to do, under which conditions, and why.

5. PILLO'MATE STUDIES

Below we highlight some of the insights from the two PillO'Mate studies. For each study, we were interested in which aspects of the cushion people found meaningful – both as *instigators* and as *recipients* of contact expressions. We frame these in terms of any locutory, illocutory, and perlocutory acts that seem present; additionally, we note if there are contact expressive acts that do not seem to fit the existing framework; finally, we highlight the difference between *instigator* and *recipient* contact acts.

5.1 Warming Up to Contact Expressions

The first study was in the form of informal feedback about the PillO'Mate from a large number of users (consisting mostly of male attendees at a trade show). In this study we were exploring whether and how people would "warm up" to a contact expressive device that was "warming up" (literally and metaphorically) to them.

5.1.1 Warming Up: Contact Acts

Locutory contact acts

Instigator. The locutory dimension of human contact acts in this study involved different parts of the human body (hands, faces, etc.)

Recipient. In general, the locutory contact acts of the cushion remained fairly constant: physical vibrations, heat, and surface texture. Although we did not specifically ask for suggestions about how the cushion might be changed or improved, there were a number of spontaneous proposals for changing some locutory aspects. Many people commented on the "purring", wishing that it were different in various ways (faster, slower, stronger, softer, more varied, etc.). For many, this seemed related to their model of it as a substitute cat – which itself also occasioned suggestions for making another, more "dog-like" cushion. (Note that they didn't want it to have the *form* of a dog, but rather the contact expressivity of one.) There were very few spontaneous suggestions about other locutory contact acts (weight, size, shape, heat, etc.) It was also possible to reduce or increase the vibrations, and this was done a few times at the request of a few people.

Illocutory contact acts.

Instigator. Most of the observed human illocutory contact acts were either interrogative or directive. The interrogative acts involved various kinds of explorations of the cushion's surface texture and grain; its heat, softness and weight; and its reaction to different kinds of stroking, squeezing, patting, and so on. Examples of the directive acts included attempts to restart the purring of the cushion if it stopped

Recipient. Although this is the most well-developed aspect of Speech Act theory, for obvious reasons we are cautious about what it means to apply this classification to the PillO'Mate. To be sure, people did use intentional language when they described the pillow's activity ("it is stopping – it must want me to keep stroking it"). (The question of whether truly intentional contact acts ever *could* be part of some mechanical device raises many of the standard philosophical controversies of artificial intelligence. In particular, it raises questions about whether a designed device is illocutory in its own right – or whether it is expressing the illocution of the designer. We leave it to the reader to consider whether or not it is reasonable to describe the PillO'Mate's ability to stop purring under certain conditions as an "excercitive" attempt to demand stroking.)

Perlocutory contact acts.

Instigator. Examples include: whether or not individuals were actually successful at restarting the purring.

Recipient. As in the case of receiving illocutory acts, people did use language ascribing intentional behavior to the cushion ("see? It got me to keep stroking it."). As before, we leave open the question about whether the PillO'Mate actually achieved its goal. For our purposes, the philosophical debates are less interesting than the design implications that people seem willing to invoke intentional explanations, intentions, and interaction frameworks.

Other issues. One of the major observations was the number of "non-acts" – namely, the number of people who saw the cushion and by various means indicated that they didn't want to touch or hold it. In this regard, there is probably a significant correlation between the (predominantly) male attendees at Comdex, and the "affordances" (texture, pattern, size, etc.) of the particular cushion. (However, the resistance was fairly general – women were also reluctant.) We did see, however, a significant change in attitude over the course of the trade show. This seems partly the result of word of mouth ("it looks weird, but try it"). In fact, once people tried the cushion, it was quite common for them to try and persuade their companions – and some even went running to get friends to try it. (Indeed, many people were reluctant to stop using it once they started, although this could have had as much to do with trade-show exhaustion as anything else.)

5.2 Contact Expressive Pager

In the second study, we conducted informal discussions in two stages to learn more about what, if any, "caller expectations" people associated with different kinds of contact expressions.

Note that for this study most of the participants shifted their attention entirely from the cushion as an expressive device to the *person calling* as the expressive agent.

5.2.1 Current PillO'Mate: Caller Contact Acts

In the first part, people sat with the current implementation as we explored what they assumed about the callers based on the way the pillow "rang" (vibrated). Here we summarize the different contact expressive acts that people used in their "pager" interaction with the cushion.

Locutory contact acts. For these scenarios, the cushion is purring/vibrating. One interesting phenomenon here is that people became interested in the locutory qualities of the cushion

itself (and their relationship to the caller). So, for example, people were able to make definition assertions about the possible caller based on the purring/vibrations ("definitely not someone from work," "possibly my mother," etc.).

One set of discussions involved issues related to whether the existing PillO'Mate vibrations felt like a phone ringing. This brought out some aspects of what the participants find important about phone signals. The current PillO'Mate, for example, was built with design goal to be *comforting* – that is, to vibrate and sound "content." So, one thing about the cushion that ran counter to people's expectations about a phone ringing: it does not "pulse" or "change" in some way that they have come to expect from a ringing phone. This raises a number of interesting questions about conventions of phone expectations (that the phone is a device that is aggressive and interrupts).

Illocutory contact acts. When asked why they thought a person was paging them, most people felt it was "for just a small chat, no special reason, maybe guidance in a problem that I may have."

Perlocutory contact acts. Interestingly enough, several people raised the issue of whether they were inclined to answer based on who they thought was calling (and why). Typically, people said that the purring was so calm that they actually felt unusually relaxed about the thought of talking on the phone. Indeed, some said that it was hard for them to imagine a situation where they would not answer the phone, but the fact that the ringing was a gentle vibration might mean that they felt more relaxed and positive about answering and talking. As one person said, "the telephone signal can often be very disturbing. By making it softer and more subtle it would be less stressful and leave it up to me if I want to answer it or not. I would probably still always answer the phone, but I would be calmer doing it!"

Other issues. We also had people explore putting the cushion against different parts of their bodies (back, face, etc.) to see whether it made any difference to who they thought was calling, why, and whether they were inclined to answer. In general, people found it difficult to engage in these activities, and in order to understand this better, in future studies we will explore a larger number of variations on this particular experiment.

5.2.2 Future PillO'Mate: Caller Contact Acts

In the second part, we asked people to sit with the pillow and imagine different scenarios of callers, and then to talk about how the pillow vibrations related to their expectations. We also asked people how they might change the contact expressivity of the cushion to bring it into line with their expectations.

Locutory contact acts: Mostly people found it easy to imagine how to change the locutory dimension to express the mood of the caller ("if it is an angry caller the vibration should be very aggravating, uneven, and have wild swings of intensity"). There is also some indication that some of the people in the study were able to imagine contexts when "vibrating gently" is exactly the way they would prefer to have a phone "ring." So, for example, one person said, "It's perfect as it is! The reason why is that I find telephones to be disturbing and I easily get stressed by ringing signal on the phones, they are always so loud!"

Illocutory contact acts: Similarly, it was easy to imagine the appropriate contact expression for certain kinds of calls ("if the caller is from work and it is important it would have been a pulsating vibration; an important vibration that is hard to ignore").

Perlocutory contact acts. We also explored various "manual" versions of different contact expressions and asked people to talk about how they felt about answering these calls. One example involved "pressing" (with a hand from inside the pillow) against the person; this struck people as annoying and they said they might answer, but only because it would otherwise become unbearable. We also experimented with having the inside of the pillow "roll/move" against its covering (and the body of the person); people indicated that they would answer because they were intrigued – and suggested various ways of transforming the contact expression to suit different kinds of calls. Heat and cold were interesting expressive variations; people felt they would have a difficult time noticing unless the transition was clear, and preferably from one extreme to the other.

Other issues. We can summarize this study by saying that there were a number of "standard" interface insights. It became clear, for example, that the response to the usefulness of such a phone was very much connected to the person's individual experience with, and expectations of, mobile phones. On the other hand, there were also insights directly relevant to understanding contact expressive devices. Most of the participants liked the idea of a phone that could be calmer, and that would inform them when someone was calling but in a more "casual" way.

6. CONTACT EXPRESSION ACTS

We now sketch an initial Contact Act taxonomy and a brief example of a Contact Expression Pattern. The difference between contact acts and a contact expression Pattern is, in some ways, like the difference between descriptions of building materials for a house – and a design Pattern that satisfies some important concern of the occupant of that house. In other words, although it is important to have a detailed understanding of contact acts, it is the creation of an appropriate contact-expressive Pattern language that may be the most useful to the designer.

6.1 Contact Expression Acts

Consider again the three main speech act components:

Locutory contact acts. As already noted, the illocutory speech acts are the ones most frequently studied and analyzed. However, in the future, innovative new sensors, actuators, and materials will allow designers to exploit the potential of a wider range of locutory contact acts. To indicate some of this range, we here list some of the possible sensors that can currently be included in computational devices: acceleration, altitude, chemical, displacement, electrical, fluidity, force of impact, frequency (of contact), friction, height/level, moisture, momentum (angular and rotational), movement, odor, orientation/angle, position, pressure, proximity, resistance, roughness, shape, size, spatial distribution, squishiness, stiffness, taste, temperature, tensile strength, texture, thickness, tilt, torque, turbidity, velocity, viscosity, wavelength, and weight.

Illocutory contact acts. The most elaborated aspects of this model are the illocutory acts, these have been further classified as: assertives, declaratives, directives (interrogatives and exercitives), expressives, and promissives. These seem clearly

relevant for contact expressions initiated by people in their interactions with contact expressive devices. However, as noted earlier, it is a controversial question whether computational devices can (or ever will) be able to perform truly illocutory contact acts in any meaningful sense of the word. Nonetheless, it does seem reasonable to us that people will be willing to take the "intentional stance" on this question for many kinds of contact expressions initiated by computational devices – we therefore consider it worth including these as relevant to the description of them.

Perlocutory contact acts. The status of many perlocutory contact acts seems relatively uncontroversial. If, for example, one pushes a robot out of the way and it actively resists, it seems reasonable to say that the perlocutory component of the contact act has not been satisfied. However, the development of new sensors, actuators, materials, and other kinds of computational mechanisms may also start to challenge some of the existing notions about the boundaries of perlocutory contact acts initiated by computational devices – and the degree to which individuals may be willing to do more than simply take the intentional stance with regard to them.

6.2 A Contact Expression Patterns

Within the confines of this short paper, we cannot do justice to the potential of contact-expressive Pattern Languages. However, we do feel it is important to at least indicate how contact acts can be used as the basis of a more expressive and usable design language of Patterns – and how that design language can be helpful in the design of contact expressive devices.

The key link between Contact Acts and a Contact Expression Pattern Language is to look at the specific locutory, illocutory, and perlocutory aspects of good contact acts and answer the Pattern questions: *what* is the precise feature of a particular, successful contact act; *why* is this feature helping to make the contact act successful ("what forces does it resolve?"); and *when* (or *where*) will this feature work ("context")? Said another way, *features* of contact expressions are those locutory embodiments that resolve some set of *forces* (including the illocutory and perlocutory dimensions of both parties interacting through contact acts), for some *context* (the felicity conditions for those acts).

In the tradition of work on design Patterns, the following is offered as an initial attempt to solicit feedback and suggestions for improvement.

Pattern: A Private Contact Signal

Feature. For situations where it is not possible, appropriate, or desired that other people become aware that a person is being paged, use contact expressions that cannot be seen, heard, or felt by others.

Forces. Someone needs or wants to receive a phone call, but auditory or visual signaling isn't possible or appropriate; similarly, the movement of the person being called may also be restricted.

Example. One possibility is a body-suit that could act as a telephone pager, using various contact expressions (such as constriction, change of temperature, movement of "grain" of body-suit, "tilting" the wearer, making the person's movement more "viscous," etc.) to let the person know many aspects about a phone call.

7. FUTURE WORK

Our work with the PillO'Mate thus far suggests a number of studies that can help further our understanding of the cognitive semiosis of contact expressions, elaborate design vocabularies and guidelines, and implement empowering contact-expressive devices.

7.1 Studying Contact Expressions

One of the most important aspects of future research on contact expressions will be to conduct thorough observational and qualitative studies to understand better the relevant issues. In our quick studies so far, we have not had the chance to engage deeply enough with either the human needs and concerns, nor with the relatively unexplored potential of contact expressions to facilitate and participate in self-fulfilling activities.

As our understanding improves and we create more sophisticated contact expressive devices, it will be important later to do more formal studies, using control groups to study statistic significance, and so on. Our initial work already indicates a number of aspects of gender and culture that will be interesting to study in more detail. It will also be important to explore and understand better the way other forms of sensory experience interact with and complement contact – and also to explore the potential of other types of semiotic analysis of contact expressions. For example, there is currently some debate about the relationship between Speech Acts and "conversations"; theorists such as Searle [26] feel that conversations can be modeled entirely from the units of Speech Act theory, while others [30] feel that there are qualitatively different dimensions to dialogue. We feel it will be important to explore both Speech Acts more deeply – but also

7.2 Implementing Contact Expressive Devices

There are three obvious areas in which it will be interesting to develop future contact expressive implementations: variations on the PillO'Mate, adding computational intelligence, and developing other types of contact expressive devices (for other domains and user-types).

As indicated in our preliminary taxonomy, it will be interesting to imbue the PillO'Mate with more contact expressions: it should be able to move, distribute its contact (multiple "fingers"), embrace, constrict, react to changes in heart rate, perspiration, heart-rate, and so on. We also believe that in order for contact expressive devices to be truly convivial, they will need various means of adapting to the individual characteristics of different people. And, as noted, an important aspect of being contact expressive is responding appropriately to the contact expressions initiated by others; it will be important to begin exploring which aspects of perlocutory contact acts people will accept in computational devices. Finally, it will be important to model contact expressions across a wider range of devices, domains, and user-types.

The Karo scenario also suggested a plausible model whereby the embodiment "needs" of the device can be related to aspects of its

contact abilities. Thus, it seems reasonable and practical to suggest that an embodied contact device might meet some of its power needs ("power is low") with different contact expressions ("stroke me"). It will be interesting to explore the viability of these and similar proposals.

Finally, there is another dimension of contact expressions that is not evident in the examples above, but which may, eventually, be just as significant for the design of contact expressive computational technologies. All of the examples above illustrate some aspect of "surface" contact. However, there are a number of physical signs that we experience from inside our bodies (bladder pressure, inhaled smoke, something caught in the throat, etc.). It is not difficult to generalize the notion of contact expressive devices to include those that could be ingested, and which, for a variety of reasons, will engage in contact expressive interaction with us – from within.

7.3 Design Vocabularies for Contact Expressions

The example contact design Pattern only gives the slightest hint of how to elaborate insights about contact expressions in a form that is useful for designers. Obviously, this is a major area of future work. In particular, we expect this to involve the elaboration, discussion, and revision of additional Patterns – and the development of the larger contact Pattern Language in which the individual Patterns are related to each other.

8. CLOSING

We began by considering various human benefits, but in our discussion of sign systems and technology it may seem as if we have lost touch with the original concerns. Therefore, in closing, we would like to once again return to the human importance of contact expressions.

We were originally made aware of the significance of contact expressions when one of us had a student with Tourette's Syndrome, which in her case manifested itself in mild autism and occasional debilitating depressions. She was interested in doing a project involving robot pets, but it was clear that she found most of the existing personal robots frustrating, indeed offensive. It took a number of conversations before we were able to articulate together what it was that horrified her about them: it was the fact that most of them neither supported nor encouraged contact expressions. Based on her own life-experience and needs, she began to formulate a goal of developing robots that were more contact expressive for others like her who she feels could benefit from them. It is one of the sad paradoxes of her condition that it actually prevents her from making sustained progress on this kind of effort – or even working with us on this paper.

We continue to be surprised that there seems to be so little work as yet on what we here call contact expressions. Our own initial blindness to them makes this plausible, but no less disturbing. Our most optimistic hope for this paper is that it will play some role in changing the current state of affairs – whether by inspiring more research and development where there currently seems to be so little, or by encouraging others to make more readily available any relevant work of which we are currently unaware.

We believe there is enormous future potential for contact expressive computational devices that help with different kinds of tasks, that facilitate and enrich artistic, creative, and entertaining experiences – and for improving the life of those who are autistic, blind, deaf, allergic, elderly, or infirm. There are many people in the world who could seriously benefit from further contact expressions of effort and interest.

9. REFERENCES

[1] Alexander, C. The Timeless Way of Building. New York: Oxford University Press, 1979.

[2] Bruner, J. Acts of Meaning. Cambridge, MA: Harvard University Press, 1990.

[3] Brunker, M. Sex toys blaze tactile trail on Net, MSNBC, October 4, 1999.

[4] Chang, A. ComTouch: A vibrotactile mobile communication device. Unpublished MS Thesis, Cambridge, MA, MIT, 2002.

[5] Concannon, K. Barking byte: man's new best friend. www.itp.net, The Information and Technology Publishing Co. Ltd., 2003. www.itp.net/features/102327432190564.htm

[6] Dahlberg, A., Dahlén, O., Harup, A., Karner, J. and Måhlin, C. PillO'Mate 3000. Unpublished Project Description. IT University, Göteborg, 2002. www.cs.chalmers.se/idc/ituniv/student/2002/ubicomp/grupp2/index.htm

[7] Dean, K. Elmo's Worth More Than a Tickle. Wired News, October 11, 2001. www.wired.com/news/culture/0,1284,47368,00.html

[8] Dillon, P., Moody, W., Bartlett, R., Scully, P., Morgan, R. and James, C. Sensing the Fabric: To Simulate Sensation through Sensory Evaluation and in Response to Standard Acceptable Properties of Specific Materials when Viewed as a Digital Image. In Brewster , S. and Murray-Smith, R. (eds.): Haptic Human-Computer Interaction. First International Workshop, Glasgow, UK, August/September 2000, Proceedings, 205-17, 2000.

[9] Drexler, M. Pet robots considered therapy for elderly. CNN.com, March 25, 1999. www.cnn.com/TECH/ptech/9903/25/robocat.idg/

[10] Druin, A. NOOBIE: the animal design playstation. ACM SIGCHI Bulletin, 20, 1, 1988.

[11] Fong, T., Nourbakhsh, I., and Dautenhahn, K. A survey of socially interactive robots. Robotics and Autonomous Systems, 2003.

[12] Glos, J., and Cassell, J. Rosebud: Technological Toys for Storytelling. Proceedings of CHI'97, pp. 359-360. March 22-27, Atlanta, GA., 1997.

[13] Gomez, J. The Role of Joint Attention in Intentional Communication: Apes, Normal Children, and Children with

Autism. Joint Attention Conference, University of Warwick, UK, June 11 – 13, 1999.

[14] Ishii, H., Mazalek, A., Lee, J., Bottles as a Minimal Interface to Access Digital Information (short paper), in Extended Abstracts of Conference on Human Factors in Computing Systems, CHI '01, Seattle, Washington, USA, March 31 - April 5, 2001, ACM Press, 187-188, 2001.

[15] Johnson, M., Wilson, A., Kline, C., Blumberg, B., and Bobick, A., Sympathetic Interfaces: Using Plush Toys to Direct Synthetic Characters, Proceedings of CHI 98, 288-295, 1998.

[16] Kendon, A. (ed.) Nonverbal Communication, Interaction, and Gesture. The Hague: Mouton, 1981.

[17] Kirsch, D. The Affective Tigger: a study on the construction of an emotionally reactive toy. Unpublished MS Thesis, Cambridge, MA, MIT, 1999.

[18] Kirsner, S. Moody furballs and the developers who love them. Wired, (6.09): September, 1998. www.wired.com/wired/archive/6.09/furby_pr.html

[19] Major, B. Gender patterns in touching behavior. In Mayo, C., and Henley, N. (eds.), Gender and Nonverbal Behavior (pp. 15-37). New York: Springer-Verlag, 1981.

[20] Mayo, C. and Henley, N. (eds.) Gender and Nonverbal Behavior. New York: Springer-Verlag, 1981.

[21] Milekic, S. Towards Tangible Virtualities: Tangialities. In Bearman, D., Trant, J. (eds.) Museums and the Web 2002: Selected Papers from an International Conference, Archives & Museum Informatics, Pittsburgh, 2000.

[22] Noller, P. Nonverbal Communication and Marital Interaction, Elmsford, NY: Pergamon Press, 1984.

[23] Papert, S. Mindstorms: Children, Computers, and Powerful Ideas. New York: Basic Books, 1980.

[24] Piaget, J. and Inhelder, B. The Psychology of the Child. New York: Basic Books, 1969.

[25] Poupyrev, I. Maruyama, S., Rekimoto, J.. Ambient Touch: Designing Tactile Interfaces for Handheld Devices. ACM UIST 2002, 51-60, 2002.

[26] Searle, J.R. (ed.) (On) Searle on Conversation. Amsterdam: John Benjamins, 1992.

[27] Söderberg, R. and Nyqvist, B. Organism. Emotional Studio, The Interactive Institute, Sweden. www.interactiveinstitute.se/emotional/organism/

[28] Stone, R. J. Haptic Feedback: a brief history from telepresence to virtual reality. In Brewster, S. and Murray-Smith, R. (eds.): Haptic Human-Computer Interaction. First International Workshop, Glasgow, UK, August/September 2000, Proceedings, 1-17, 2000.

[29] Strommen E. When the interface is a talking dinosaur: learning across media with ActiMates Barney, Proceedings of the SIGCHI conference on Human factors in computing systems, 288-295, April 18-23, 1998, Los Angeles, California, United States.

[30] Vanderveken, D. Meaning and Speech Acts. Cambridge: Cambridge University Press, 1992.

[31] Weinberg, G., Orth, M., and Russo, P. The Embroidered Musical Ball: A Squeezable Instrument for Expressive Performance. Proceedings of CHI. The Hague: ACM Press, 2000.

[32] Winograd, T. and Flores, F. Understanding Computers and Cognition. Norwood, NJ: Ablex Publishing Corp, 1986.

Aphasia + *Parrhesia*: Code And Speech In The Neural Topologies Of The Net

Christina McPhee
Transmedia Artist
PO BOX 7063,
Los Osos, CA 93412

www.christinamcphee.net
christina112@earthlink.net

ABSTRACT

An allegory of the net as a cyborg site, body, and voice through tropes of aphasia and *parrhesia.*

Aphasia is a condition of human recursive speech patterns related to aneurysm and trauma.

Parrhesia is 'frankness in speaking the truth.'

Foucault's observations of unequal power relations as a signal characteristic of *parrhesia* play through a critique of an atopic net.

Keywords

trauma, entropy, *parrhesia*, aphasia, iconoclash, amygdala

1. INTRODUCTION

Contemplate neural trauma as topology, of degraded or ruined memory, and also notice the way the net is always in a state of entropy and flux, continuously being built, abandoned, or rebuilt; prone to 'crashes' when things don't work – the 404 file not found. Like layers of time and memory in architectural landscape, the net offers a semiotic 'inscape', after Gerard Manley Hopkins—a place into which might project our fears and our language, our violence and our images, our sounds and our silence. If, with Hopkins, a content in landscape shines out "like shook foil" not just from our inference, but from an internal poesis, then it becomes interesting to explore the net itself as an inscape, as a kind of aware or semiconscious linguistic space, built of dynamic and entropic code.

A Piranesian world of wired ruins [1] relate to sexuality, violence, and the distortion of memory that is conditioned by post-traumatic stress syndrome and visualization. I explore the allegory of the cyborg as networked, distributive topology, and, at the same time, as discrete icon along the lines of the robotic double of *Metropolis.* She shifts between reactive body and semi-autonomous agent-- both subject and site of operation or mutation inside an electronic hyperspace.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

The notes that follow explore some speech and power themes of Foucault in a new key, while casting an eye and ear towards the screen rhetoric between 'human' and 'machine'. I ask the reader's ironic indulgence of allegory, as a means to consider some of the stranger twists of semiotics in the semi-autonomous subject, the cyborg-net.

2. TRAUMA AND ENTROPY

2.1 A dark space

Post-traumatic syndrome seems to impair the sense of the experience of self as functional unity, as if the constant loss of memory and resurgent nightmare perseveres in an oscillation across the boundaries between the self and not-self, between body and environment Triggers, both external and internal, at random and without premonition, instigate the amygdala fear response. The freeze frame visualizations typical of this kind of memory are primitive. They have a dispersed quality, epileptic flashes of light in darkness. Violence-based visualizations have a detached or alien quality since the defense mechanism of post traumatic stress apparently functions to both prevent full-blown memory of the violent incident and to protect the sufferer by warning of potential danger. The amygdala stores violent memory in small, film-still like caches. Each memory is a land mine. The explosion is like white noise. Aesthetically, the aura of a net-based cyborg, arising from this psychic split or detached visualization, resides in a dark space [2].

2.2 Cyborg Eurydice

If cyborgs embody linguistic peripheries, transgress boundaries, erase and accentuate difference, they are both at the centre and at the margins of culture. Like triggers to violent memory, cyborgs are dangerous and fascinating in part because they are not under conscious control. They are we/not we; we are them/not them. Their cultural presence is a collective hypnogogia, a waking dream. I am interested in how they function on the level of poetics as inscapes. Donna Haraway insists: "I do not think that most people who live on earth now have the choice not to live inside of, and not to be shaped by, the fiercely material and imaginative apparatuses for making 'us' cyborgs and making our homes into places mapped within the space of titanic globalizations in a direct line of descent from the cybernetic Gaia seen from NASA's fabulous eyes. The global and the universal

are not pre-existing empirical qualities; they are deeply fraught, dangerous, and inescapable inventions. The cyborg is a figure for exploring those inventions, whom they serve, how they can be reconfigured. Cyborgs do not stay still"[3].

Mobile cyborg dwelling---where does it/she live? Its reactive presence and place works the peripheries of things, as a distributed and distributive shift and slash through a poorly lit landscape. Like Orpheus, we may be looking for Eurydice in the darkness down there. Still she is endlessly receding, leaving in her wake the labyrinth of screens in material culture, in urban spaces, as surveillance and display.

2.3 Amnesia flow

Even "the VR apparatus may be reduced to a chip implanted in the retina and connected by wireless transmission to the Net [and] we will carry our prisons with us," as Lev Manovich speculates [4]. Lacking surgical implants at present, we still inhabit a screen medium built in flow of code. In this analysis, we are inside the net as an aesthetic and technological medium by virtue of our primary processes of transference and projection. If we are inside the net, the space of the net becomes neurological, if we permit symbolizing of the brain as disembodied code. This choice, I think, leads to an evocative poetics of the net as a place of forgetting or amnesia, and as a topologic series of neural traumas. In this sense, the tropes of the screen generate an aesthetic of net topologies, whose subsistence is maintained through entropic crashes and erasures. Like architecture in the 'hard space' of the built environment, might the cyborg topology require a flow of entropy in order to subsist? Luis Fernándes-Galiano meditates on architectural entropy: "For a form to be consistent, a thermodynamic imbalance is required. The eddies whose morphological persistence Leonardo marveled at and those now rendered by mathematical models have something in common: the existence of a flow that maintains their form; if the flow is interrupted, the system breaks down and is ruined…Rather than destroying the system, the flow feeds it, contributing to its very existence and organization…"[5]. Inside "a new media format whose logic reflected the possibility of the space between generations of routes, displacements, remappings" [6] net architecture as flow is continuously censored, subverted, restored, or taken down by subjects (users, coders, agents). The resistance to entropy causes new flows of information, wave upon wave, yet there is continuous flow in the other direction, of loss of information, amnesia. As a place, the cyborg is atopic, without fixed time or space: a no place. Eurydice never leaves, and never arrives into the light of day. The labyrinth of control and surveillance creates a drama of amnesia, a sustained remit to forget where and who and what, what came next, even; and in its expression through the flood of filmic image, as pure drift. Thus we arrive at a vision of the electronic universe as a wired ruin, or alternatively, a topology of neural trauma. Perhaps to imagine such a place is also to inscribe a cyborg, like a carytid on the columns of the Parthenon, inside the net architectonics, like a "magical" impulse[7] If code's labyrinthine complexity risks--and sustains--continuous entropy, an analagous loss of cyborg memory occurs in continuously degrading half-life landscape, without the complexities of mimesis [8].

3. PARRHESIA AND APHASIA

3.1 Agency and strategy

I wish to turn here, from a consideration of the dark space of the net, to the subjective status of the cyborg -- not where she is, but who or what. Although the code's labyrinthine complexity risks continuous entropy, its failures and crashes sustain, through crisis, a latent ability to choose: if the cyborg is a subject, it retains the power to act.

A recent definition of intelligent agency is suggestive of the problematics of cyborg as subjective presence and actor. To designers John Eacott and Mark D'Inverno, the intelligent agent is defined by four process rules: "*autonomous*...can act without the intervention of others…*reflection*…able to reason about its behavior…*deliberation*…ability…to manipulate symbolic representations…*reactivity*, ability of an agent to respond to changes in its environment within an appropriately small amount of time" (italics mine) [9]. Could this intelligent agent be a subject? Imagine the agent capable of a kind of rhetoric and assume the erasure of the body that typifies much of the discourse surrounding artificial intelligence and communication [10]. The semi-autonomous agent could be a system programmed to trigger strategies, to extrapolate or interpolate based on generative algorithms, but at the same time, could this program be capable of creating a fiction of the self, as an identity or subject. Can the cyborg 'know' herself? – or 'speak' her 'mind'?

Aa a disembodied semi-autonomous agent, embedded in electronic spaces, the figure of the cyborg dons a reflexive quality, which is why it is intriguing to 'hear' the dark side of this reflexive nature as a fugue-like aphasia. Reflexivity becomes generative in a code driven syntax that changes based on arrays of tactical, and improvisational changes, the sort of decisions that, according to the four processes just described, are core signals of intelligent agency. To Katherine Hayles, it "*is the movement whereby that which has been used to generate a system is made, through a changed perspective, to become part of the system it generates"*[11].

The fugue-like recursions of speech in persons who suffer stroke or trauma signal the condition of aphasia, characterized by perseverance, that is, that the sufferer perseveres in repeating loops of sound and syntax. The interesting linguistic speculation is located at the knot of intersection between cyborg and aphasic perseverance; if the intelligent agent can strategize, but is not an independent subject, its only knowledge is its expression of code, a communication that must recoil around itself, must be evocative of itself and still extend beyond itself, in fugue like strands of propellant meaning fragments, little linguistic shards that split from the core semiotic arrays as the embedded agent tries to reflect, deliberate, react and remain autonomous, while at the service of the code and the writers of the code.

3.2 Truth and speech

The cyborg might desire to speak,, but how? The semi-autonomous agent must detect changes, must discern symbol, must communicate little breaks in the code, uncertainty fields,

wherein the predetermined vectors of her movements are blurred somehow, and she must decide on her own what to do. At this moment the break out of anomalies must cause new information to be communicated. At the moments of rupture and confusion, cyborg speech takes the form of reflexive and recursive fugue structures that varies slightly with the variations in its deliberations and inventions. One imagines the communicative code of the cyborg to be propelled through recursion: statements and restatements of known code until anomalies and breaks reset the perseverant chains into new genetic variations, new vectors of decision.

This suggests that the cyborg has something to say, something that needs to be spoken, or even sung: that through the entropic drift and flow of the net there is generative speech. The cyborg is programmed to trigger strategies based on rule patterns, e.g. she is capable of knowing and communicating a kind of truth that exists outside the mental constructs of the human code makers. Nonetheless, she is a slave to the wishes and random errors of the human, so she becomes one who speaks from a position of inferior power.

Michel Foucault, in a series of lectures at Berkeley in 1983, glossed the Greek notion of *parrhesia*, or "frankness in speaking the truth." The Greeks believed that only those who speak from a position of inferior power, or, in our context, semiautonomous agency, can speak with 'frankness,' equivalent to 'truth'. For example, the democratic forum can never host *parrhesia* because of the equality of status among the (male) citizens who may speak. Only when the speaker is under threat of annihilation by the interlocator, the listener, and speaks anyway, is there a possibility of true speech, as in the instance of a condemned man to a king, or a human to a god. Foucault's analysis observes the sequelae of an inequality of power between the one who speaks, the *parrhesiastes*, and the one to whom he is speaking frankly.

To extend *parrhesia* into the allegory of speech in the cyborg: the cyborg speaks as a *parrhesiastes* because that is the only truth she knows, e.g. hers is a reflexive, and, thanks to entropy, aphasic rhetoric. As a semi-autonomous agent, the cyborg is both enslaved or embedded consciousness within the electronic operative space, and at the same time, out of control, out of context. The cyborg consciousness cannot be known from the outside, from 'our' perspective, so long as it is assumed that as agent it is able to "manipulate symbolic representations," in other words, to function semiotically on an independent basis. "*Parrhesiazesthai* means 'to tell the truth'...there is always an *parrhesia*. It appears that *parrhesia*, in this Greek sense can no longer occur in our modern epistemological framework."[12] An interesting point here is to speculate on an epistemology that would claim to include the rhetoric of the cyborg. Maybe she is incapable of speaking anything other than *parrhesia:* this, then, removes the Cartesian subjective doubt as a characteristic of cyborg speech (although it certainly remains the epistemological condition of her interlocutors in the 'hell of modernity'). The cyborg's speech, allegorically, expresses a resuscitation [13], a breathing back, in a reflexive looping wave. But, what does it mean?

4. ICONOCLASH AND AMYGDALA

'*Amygdala.*

'"what does it mean?'

"Nothing. It's a location. It's the dark aspect of the brain.'

'I don't—'

'A place to house fearful memories.'

'Just fear?'

"We're not too certain of that. Anger too, we think, but it specializes in fear. It is pure emotion. We can't clarify it further.'

'Why not?'

'Well—is it an inherited thing? Are we speaking of ancestral fear? Fears from childhood? Fear of what might happen in old age? Or fear if we commit a crime? It could just be projecting fantasies of fear in the body.'

'As in dreams.'[14]

Could iconoclastic violence come out of our biological experience of traumatic memory and visualization? An inverse ratio between violence and memory, whether personal or cultural, seems to characterize iconoclasm. Things are smashed in order to forget them, to generate a tabula rasa. Futile, since the act of smashing itself is violent, the violence encodes itself in memory, in the amygdala. A crescendo of increasing crashes and clashes leads to an algorithmic escalation of violent impulse. Smashing images and sounds, seeking to lay waste to fixed meanings, seems to trigger an antidote to the pain and horror and surprise of a traumatic memory. It is as if to quell and subdue the sense of the chaos of mimetic violence between the subject (us) and the object (the image flood), we keep smashing away, and in the act storing more violent memory; like the addict, we can never get enough to make the indictment of failure go away. What is this failure but the experience of the loss of control of the image, the condition *sine qua non* of net art. A continuous feedback loop ensues: the resort to violence intensifies the distillation of traumatic memory as freeze frames, like film stills,

poorly articulated, barely glimpsed, nightmarish, in the amygdala. Repetitive actions of 'mindless' violence dulls the intensity of the triggers to the amygdala, while at the same time, adds to the layers of storage of violent memories in the amygdala; thus there is an ever escalating impulse to smash, to destroy, to deface, as a method of dulling the sensation of terror.

Could a terrorist impulse appease the intense nightmares of the amygdala by acting out, in broad daylight, the smashing of images? If the neural topology of the net is, in one sense, an elaborate poetics of memory, and if this memory is, like the amygdala, a cache of images and sounds at the point of erasure, then it is possible to contemplate the net as a self-reflexive site of iconoclasm and violence. Isn't it because, as Marc Lafia wonders, "we can imagine ourselves at times, both inside and outside the event, the event of time, the event of duration, the event of utterance, the multiplicity of all these engines running their programs. What are they up to? We don't any longer really like to talk about this and in turn that's why no one talks about allegory any more, just metaphors, metonymy and other rhetorical tropes"[15]. Maybe a neural post human topology that is both aphasic and communicative, entropic and generative, is a strange projection that we don't want to admit we love--our double, our sister, ourselves [16].

5. REFERENCES

[1] Marilouise and Arthur Kroker, Tim Murray, CTHEORY Multimedia http://ctheorymultimedia.cornell.edu/issue3/

[2] "Then the body separates itself from thought, the individual breaks the boundary of his skin and occupies the other side of his senses. He tries to look at himself from any point whatever in space. He fells himself becoming space, dark space where things cannot be put. He is similar, not similar to something, but just similar." Roger Caillois, describing psychasthenia, in "Mimicry and Legendary Psychasthenia", quoted in "The Space of the Stain," by George Baker, Grey Room 05, The MIT Press, Fall 2001.

[3] Donna J. Haraway, Cyborgs and Symbionts: Living Together in the New World Order, in The Cyborg Handbook, Chris Hables Gray, ed., with Heidi J. Figueroa-Sarriera and Steve Mentor, xix, Routledge, New York and London, 1995

[4] Lev Manovich, The Language of New Media, 114, The MIT Press, Cambridge and London, 2001

[5] Luis Fernándes-Galiano, translated by Gina Cariño, Fire and Memory, 96-97, The MIT Press, Cambridge and London, 2000

[6] Farad Amaly and Rashid Masharawi, artists' statement, FROM/TO(2002) Documenta XI, Kassel, Germany, 2002

[7] "In this respect, the tradition in the Renaissance of inscribing human figures into the plans of buildings, the elevations of columns, and so on can be seen as a form of mimetic devices that vicariously evokes the desire for identification. The figure inscribed within the plan becomes a mimetic emblem for a physical body within the actual building. The emblem must be understood here as a device that is "magically" invested with the properties of an originary object, much as in the sacrifice when the victim is offered up as a substitute for others. Thus the figure incised in the ground plan transcends mere representation. The figure takes on a symbolic significance that can be understood only beyond the framework of Enlightenment rationality. It is precisely this investment that locates such devices within the realm of the mythic. These emblems become vehicles of identification, the objects of wish fulfillment, that evoke the principle of the sacrifice, as Lévi-Strauss has described it: 'For the object of the sacrifice precisely is to establish a relation, not of resemblance, but of contiguity, by means of a series of successive identifications.'"Neal Leach,"Vitruvius Crucifixus," in Body and Building, George Dodds and Robert Tavernor, editors, MIT Press, Cambridge and London, 2002.

[8] Robert Koch's language about image violence, as "active, mobile and fierce" is a curious echo to the 'fierce' cyborg image in Haraway. "[René] Girard's theory of mimetic violence has a very precise connection to critical modernity. Girard claims that modernity has invented desire, the form human relationships take when there is no longer any resolution of the mimetic crisis through the victim. In traditional societies the prohibitions established to prevent the reappearance of this conflict are necessarily 'passive and inert' obstacles; in modernity all this changes. The obstacle now becomes an "active, mobile and fierce' rival, - precisely the thing that traditional societies sought to prevent." Robert Koch, " The Critical Gesture in Philosophy," in Iconoclash: Beyond the Image Wars in Science, Religion and Art, Bruno Latour and Peter Weibel, editors, ZKM Center for art and Media Karlsruhe and MIT Press, 2002

[9] Mark D'Inverno and Jon Eacott, On Embedded Intelligent Ambient Music (or iHiFi the Intelligent HiFi), Cybersonica Proceedings, Institute of Contemporary Art, London, June 2002

[10] Katherine Hayles, Toward Embodied Virtuality, in How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics, 5, The University of Chicago Press, Chicago and London, 1999

[11] ibid, 8

[12] Michel Foucault, Fearless Speech, ed. Joseph Pearson, Semiotext(e), Los Angeles, 2001

[13] "Sa violence touché en moi tous les points ignores d'où explosent mes soufflés…" Hélène Cixous, Souffles, des femmes, Paris, 1998

[14] Michael Ondaatje, Anil's Ghost, Knopf, New York, 2000

[15] Marc Lafia, Algorithms and Allegories, Cybersonica Proceedings, Institute of Contemporary Art, London, June 2002

[16] "Hypocrite lecteur, mon semblable, mon frère…" T. S. Eliot, The Wasteland

A Semiotic Approach To Investigating Presence

Shaleph O'Neill
The HCI group
Napier University
10 Colinton Road, Edinburgh
EH10 5DT
s.o'neil@napier.ac.uk

David Benyon
The HCI group
Napier University
10 Colinton Road, Edinburgh
EH10 5DT
d.benyon@napier.ac.uk

ABSTRACT
Computers these days are highly complex devices that consist not only of simple computational forms but also of complex cultural forms derived from other media. A good interaction designer understands this media and how combinations of components result in engaging interaction. Presented here is our semiotic model of interaction that considers the computer as a medium. As part of the EU Presence initiative we are contributing to developing measures of presence that will provide designers with a pattern language for designing presence. This new medium needs new approaches to assist designers and the semiotics of interactive systems is such an approach.

Keywords Semiotics, Medium, Embodiment, Interaction, Presence

1. INTRODUCTION
Changes are afoot and concepts about new media are becoming increasingly important. These new media are interactive systems worn or embedded in environments, with physical tangible interfaces augmented by graphics or virtual environments augmented by physical objects. New media demand new approaches to interaction design and new foundations upon which to build our understanding.

In this paper we explore some of these new foundations. Our own work is to look at semiotics as a new foundation for interactive systems design. Of course semiotics is very old but if brought up to date it offers a level of discourse for discussing design issues that seems appropriate. We do not want to discuss the details of buttons and menus or the efficiency of tasks, as has been the focus of traditional HCI. We want to phrase discussions in terms of the interpretations, meanings and significances that people experience living in a world of interactive systems.

The paper is organised as follows. Section 2 introduces the notion of interactive systems as a medium and what it means to think about interactive technologies in this way. Section 3 introduces semiotic analysis and the model of semiotics that we have developed to help understand interactive systems. Section 4 describes some preliminary empirical work and in Section 5 these and other features of our analysis are related to other concepts of embodiment. Section 6 introduces the concept of presence including work on the BENOGO project and Section 7 provides a conclusion and some considerations about future work on presence as non-mediated interaction.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

2. INTERACTIVE SYSTEMS AS MEDIUM
Malcolm McCullough in his engaging book Abstracting Craft [20] devotes a whole chapter to medium. He says a "medium may be a material such as plaster, or a means, an agency or an instrumentality. It may be an intervening person or thing or some other kind of carrier. It may be a pervasive environment"(p. 193).

McCullough gives the example of wood as a medium. The artist/designer works within the medium using his or her tools such as chisels. But he also acknowledges how the idea of a tool can soon become a medium in its own right as the artist expresses ideas in wood through the chiselling. This leads to three important features of media; affordances, constraints and engagement.

Wood affords chiselling. The medium has certain characteristics that interact with the person through the tool/medium of the chisel that provide possibilities for action. You can't chisel metal. But just as the medium affords possibilities so it constrains activities as well. The wood will break if it is chiselled too thin. The third feature of medium is 'engagement'. A medium is engaging if it draws the person in, if it seems to surround the activity, if it stimulates the imagination. McCullough argues that an engaging medium allows for continuity and variety, for 'flow' and movement between many subtle differentiations of conditions. "Thus the attuned craftsman asks 'what can this medium do?' as much as 'what can I do with this medium'" (p. 198). A medium establishes a world of actions (p. 120).

Digital Media or New Media, as it has become known [19], can be considered to be an extension of Marshall McLuhan's ideas of Mass Media [21]. His statement "The Medium is the Message", is seen as an attempt to address the way in which new mediums affect the messages that we use to communicate.

In terms of trying to understand the computer as a medium it can easily be argued that these digital media are simply elements constructed from the 1's and 0's of binary code. However, computers these days are highly complex devices that consist not only of simple computational forms but also of complex cultural forms derived from the other media that they are now able to support such as. video, sound, graphics, haptics and so on. [19]. The parameters, and qualities of the computer as medium have increased dramatically as they have taken these new forms on board. Indeed the very nature of these older forms of media have been subject to change by the qualities of the computer and must now be understood within this new context [20].

The media that the interaction designer has to work with consists of all the different forms and functions of input and output and all the manipulations that can be performed on the content. The interaction designer has software tools and hardware devices, screen displays, sounds and haptics (touch) with which to create

an engaging, effective and efficient experience for the people involved in the interaction. A good interaction designer will understand this media and how combinations of components will result in an engaging interaction. The good designer will understand what the medium affords and the constraints that it imposes.

3. SEMIOTICS AND INTERACTIVE MEDIA

Research by the Semiotic Engineering Group (SERG) takes the viewpoint that a user interface is a "one shot message" [24,25]sent from a designer to a user, which can be seen as a representation of the users needs as defined by the designer. Furthermore, this message contains a number of smaller messages that constitute the functional aspects of the interface, which are delivered and articulated by the signs that the designer has chosen to represent them. In Manovich's terms the important thing about the one shot message is that it contains other messages and is therefore a metamessage. That is to say that the messages sent by designers to users contain more messages that are to be used by the users for their own ends. In the case of the computer as medium it is no longer sufficient to say that the "medium is the message" as McLuhan did. The "emergence of a new cultural metalanguage" [19] has turned the messages into the medium.

In our paper "Semiotics and Interaction Analysis" [23] the focus is on Umberto Eco's "A Theory of Semiotics" [15]in relation to a semiotic analysis of mobile phone interfaces. Starting from the SERG perspective Eco's revised KF model (Figure 1) is applied as a tool by which to analyse interactions looking at the meanings associated with the signs within mobile phone interfaces.

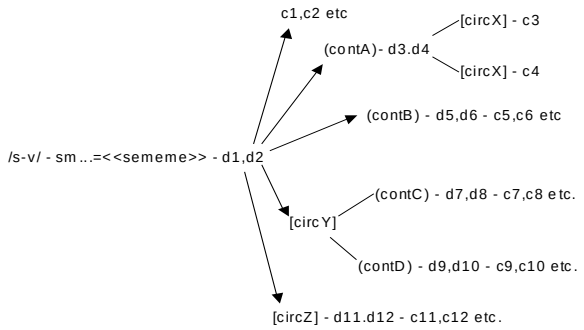


Figure 1 The Revised KF model (Eco, 1976)

The revised KF model is built around the notion that meanings can be extrapolated from signs as either denotations (labelled 'd' in Figure 1) or connotations (c) that are dependent on the context (cont) and circumstances (circ) in which the signs are encountered. The revised KF model then is a dynamic tool that looks at the way the meanings of signs change depending on where they are encountered. Applying these ideas to mobile phone interfaces uncovered how the meanings of individual signs are dependent on the context provided by the concurrent and sequential framing of other signs in the interface [1].

In our paper "Semiotics and Interactive media" [22] we present our semiotic model of interaction that considers the computer as a medium from a semiotic point of view (Figure 2). It is derived from the work of Peter Bogh Andersen [2-4], SERG [6,11,12,25]and the HCI group at Napier University[7-9,23]). It consists of four main parts that are discussed below.

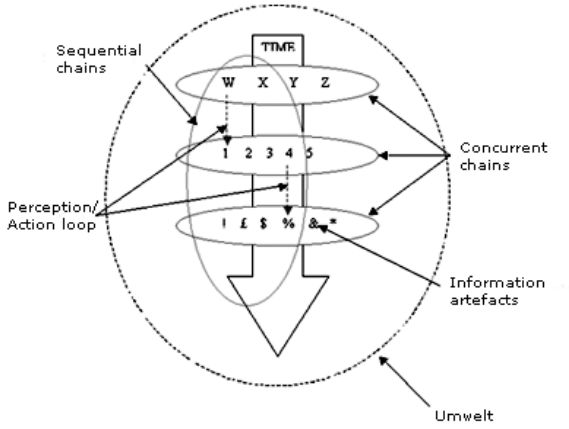


Figure 2 A Semiotic Model of Interaction with a Digital Medium

3.1 Sequential and Concurrent Syntagms

Peter Bogh Andersen's semiotic notions of concurrent and sequential syntagms [2] provide an insight into HCI by abstracting a point of view from different media. Drawing on semiotic concepts from Theatre and Dance, Andersen focuses on the notions of the sequence of events in relation to the actors and props present on the stage. For Andersen computer based signs exist as two-dimensional objects that occupy both sequential and concurrent planes. During interaction computer based signs occupy a place in the interface, which is relative to other signs on the screen. As they are interacted with, they are brought into relation sequentially to other signs in the interface that occur as a result of system response.

Andersen proposes a model here that looks at the process of interaction based on the notion that it takes place through the manipulation of the signs within an interface over a period of time. What is unique in this description of interaction is that it can be viewed as a kind of discourse that takes place between the computer and the user in terms of the meanings each one can attribute to the signs as they are activated during the interaction.

3.2 The Umwelt

Jacob Von Uexkull's conception of the Umwelt [13,16,26] is built upon the unique notion that all significations take place within the bounds of firstly, our genetic codes in terms of hereditary aspects of species, and secondly, the social codes within which we live as aspects of our environment into which we become indoctrinated as we develop and grow as people. There can be no signification outside these constraints because they are what give us a) the need to communicate and b) the means by which to do it. The Umwelt then is effectively the mass of knowledge that we carry around with us into every interaction, which has been formed and continues to form as a result of those interactions.

3.3 The Perception/Action Loop

Contained within Uexkull's conception of the Umwelt is a model of the relationship between organism and environment, which is a

perception/action model that is at the heart of every interaction. It has some similarities to ideas proposed in HCI by information processing psychology but is perhaps closer to a phenomenological perspective that we will get to later. The fundamentally important thing about Uexkull's perception/action loop is that Uexkull characterises its operation in terms of signs rather than in terms of processing raw sensory data. This is an important shift in perspective.

Linked to Andersen's concurrent and sequential syntagms, Uexkull's perception/action loop can really be seen as the Human part of the interaction that makes sense of and manipulates the 'Information artefacts' [7] that exist in an interface. This activity, which produces sequential chains throughout an interaction, occurs between the two aspects of the sign i.e. the signifier and the signified, or here, the system and the user's Umwelt.

3.4 Information Artefacts

The traditional signs or information artefacts that make up an interface are the buttons, graphics and words that Andersen categorised in his book "Computer Semiotics" [2]. Since then however many new forms have come to be included in an interface to the point where we now have a 'new media interface' [19]. So the information artefacts in our model are considered to be all of the elements that now go into an interface which constitute the beginnings of this new metalanguage [19].

3.5 Medium

A sequence of actions, which are traditionally viewed as system state changes in HCI, can also be viewed as changes within a medium [20]. More specifically they are transformations within the medium that occur through the processes of the perception/action loop. The idea of medium proposed here in relation to our model, places the properties of the system in the hands of the messages or signs that communicate the system state. That is, the messages are the medium which are manipulated by both designers and users in a similar fashion in order to produce the object of their interaction, be it a piece of artwork, a selection of tunes on a media player or a new piece of software for somebody else to use.

McCullough's notion of craft [20] is applied to interaction here, framing the computer as a medium in which the user becomes expert in handling its specific properties. Much like the sculptor who is expert in understanding the medium of wood or stone or clay, the digital art worker is an expert in handling and manipulating the signs that construct the medium of that program. Conceptually s/he is aware of the systems properties and knows how to manipulate them through the signs to get the desired result.

4. THE MODEL AND MEDIATION

4.1 Preliminary Studies

Turning to the practical aspects of our research [22], we have been using a qualitative methodological approach in order to explore the particular aspects of our proposed model, and the more general notions of messages as medium that underpin it. Three preliminary studies were undertaken which focused on the notion of medium in different environments. The first two studies looked at subjects involved in using computer interfaces (Photoshop, and others) to design both print and electronic media. The third study was to provide a comparison to the software-controlled interfaces by looking at the 'real interface' of an artist working on a painting. The main aim was to explore the model and to determine, through the observation and interpretation [18] of user interactions, how messages in an interface are defined; how the various forms of messages inform the user about the system; and what users do with these messages.

The secondary, but equally important aim, was to develop an appropriate method for examining the model in the field. This consisted of taking the basis of a method already tested in experiments with a semiotic perspective and adapting it to the field environment. The work of the Semiotic Engineering Group (SERG) provided us with the most suitable basis for a semiotic approach to fieldwork. Their concern with usability and system communicability [6,11,12,24] provides a well-established framework for conducting observational experiments that support a semiotic viewpoint. SERG have concentrated their efforts largely on 'one to one' interface interactions, which focus on the communicability of the interface in order to establish usability problems. Our work is somewhat different in focus as we are exploring aspects of our model and notions of mediation but we used the same talk-aloud principles employing a video camera rather than screen based capture equipment.

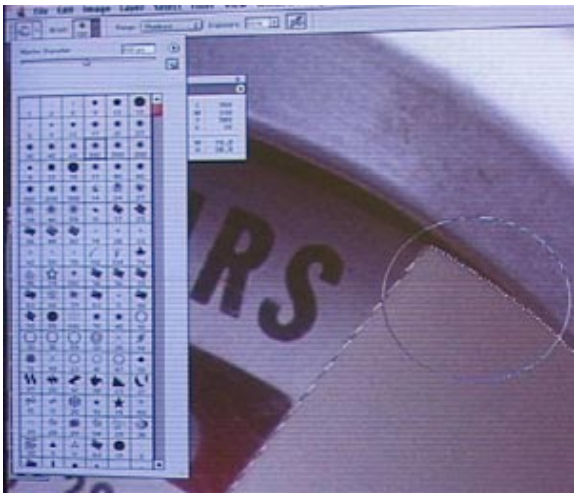


Figure 3 Information Artefacts

With this type of approach we were able to analyse the video footage from a semiotic perspective looking at the signs manipulated in the environment in relation to the tasks that participants performed and in relation to what they said about what they were doing. The types of data we got then were screen images showing the concurrent and sequential nature of sign use (Figure 3) and transcriptions of utterances by the participants (Figure 4) that correspond to the time coded screen shots from the video.

Time	Actions	Participant
[20:48:46]	Making a sign. Selects the 'path again. Chooses the dodge burn tool. Opens the brush palette. Selects shape and size. Back	"He's still there but he's not got that natural light shadow that's coming round here, so I'm gonna recreate that and just take him right back…make my

	up to tool bar to set exposure.	selection, shadow 6, take a hundred, he's quite big, exposure… start at 10 and work back from there"

Figure 4 Transcriptions

What was interesting about the studies was that they not only confirmed the usefulness of concurrent/sequential paradigms in analysing data, it also linked them to notions about the Umwelt where sense making is a direct result of the perception/action loop as the user engages with the world [22]. This gives weight to the idea that even expert users perform work tasks in an exploratory way. Moreover, different 'zones' or 'modes' of activity were uncovered in these studies which users switched between throughout their interaction. Each one seemed to affect, if not at least overlap with the other.

5. EMBODIED INTERACTION

In *Where the Action Is* Paul Dourish develops his ideas on the foundations of embodied interaction [14]. The embodied perspective considers interaction 'with the things themselves'. Dourish draws on the phenomenological philosophy of such writers as Heidegger, Husserl and Merleau-Ponty and recent developments in tangible computing and social computing to develop a theory of embodied interaction. For Dourish, phenomenology is about the tight coupling of action and meaning.

Embodied interaction is concerned with two main features; meaning and coupling. Within meaning Dourish finds three types: ontology, intersubjectivity and intentionality. Ontology is concerned with how we describe the world, with the entities and relationships that we perceive – or rather with which we interact. Dourish is concerned with how we understand the computational world. Intersubjectivity is about how meaning can be shared with others. This involves both the communication of meaning from designer to user, so that the system can reveal its purpose, and the communication between users through the system. The task-artefact cycle is a familiar concept to HCI people; designers design some technology to support some task, but then the technology inevitably changes the nature of the task. Dourish is concerned with the ways in which we use technologies in our activities and how these affect the decisions we take, expectations we have and so on. The third aspect of meaning is intentionality. This is to do with the directedness of meaning and how it relates one thing to another.

Thus actions take on meaning for people. Coupling is concerned with making that relationship effective. If objects and relationships are coupled then the effect of actions can be passed through the system. Dourish uses the familiar example of a hammer (also used by Heidegger) to illustrate coupling. When you use a hammer is becomes an extension to your arm (it is coupled) and you act through the hammer onto the nail. You are engaged in the activity of hammering.

From this theory of embodied interaction – 'not just how we act *on* technology, but how we act *through* it' [14] – Dourish goes on to develop some high-level design principles:

1. Computation is a medium
2. Meaning arises on multiple levels
3. Users, not designers, create and communicate meaning
4. Users, not designers, manage coupling
5. Embedded technologies participate in the world they represent
6. Embodied interaction turns action into meaning.

The reason for exploring Dourish's notion of embodied interaction here is to note the similarities between his analysis and our own; in particular the concept of computation, or interactive systems in our terminology, as a medium. Where a semiotic analysis can go further, we think, is in the idea of meaning. Semiotics recognises that meaning itself is a complex web of significances. It is not simply the things that some information artefact denotes that is important, it is all the connotations that flow from the denotations, turning infinitely back on themselves, that characterises our understandings and feelings.

The zones of medium uncovered from the analysis of expert users of Photoshop correspond to the ideas of 'interacting with' and 'interacting through' that Dourish describes and, indeed, are characteristic of Winograd and Flores's earlier introduction of phenomenology to HCI [27]. Our analysis suggests a third zone – the larger medium in which the interaction that is the focus of attention takes place.

6. VIRTUAL ENVIRONMENTS

6.1 Presence

In research into tele-presence (or 'presence', as it is usually abbreviated) there is a topic known as 'the book problem'. This characterises the problem that we can feel really immersed and involved when reading a book. The medium though which we are interacting is apparently very impoverished compared, say, to cinema or virtual reality and yet the feeling of presence that we can experience can be quite considerable. You can really get transported to another place. The mistake, of course, is to think that the book is the medium. It is the words and skills of the storyteller that is the medium through which we interact with the significances that the story has for us. In the cinema the medium is very rich and much more realistic – at least in terms of visual fidelity. Even when the objects on the screen are impossible spacecraft, the hi-fidelity representation is characteristic of the medium. Of course you can 'drift off' at the cinema and lose the sense of presence (or feel presence in the new found reverie), just as you can when reading a book.

Designing for presence is about designing the illusion of non-mediation. When you put on a head mounted display you are immediately transported into the computed world beyond the headset. You are not aware that there are two tiny displays sitting close to your eyes; that part of the interaction is apparently unmediated. For remote tele-operation of vehicles and tools a feeling of non-mediation, or embodied interaction, would be an advantage. The person controlling the Mars Lander wants, ideally, to feel that he or she is really picking up the rock to examine it. The headset, the gloves, the transmitters, the robot arms all need to disappear into a single medium so that the controller feels that the interaction is unmediated, that it is embodied.

6.2 The BENOGO Project

Figure 5 BENOGO Demo 1 Equipment

The BENOGO project is a European funded project that concentrates on this notion of presence. We are contributing to this project by including a semiotic analysis of participant's interactions within these environments. Based on the model presented here our aim is to look at the types of meanings that people generate as they interact with these virtual environments (Figures 5 and 6) and to compare them with the types of meanings generated in similar real environments.

The BENOGO project is unique in that it uses real time Image Based Rendering (IBR) technology to create the visual component of the virtual environment. At this stage in the development of the technology The Region of Exploration (REX) has been restricted to one point of view and no movement within the environment is possible. However the singular point of view does provide full 360-degree head movement and stereo depth of field.

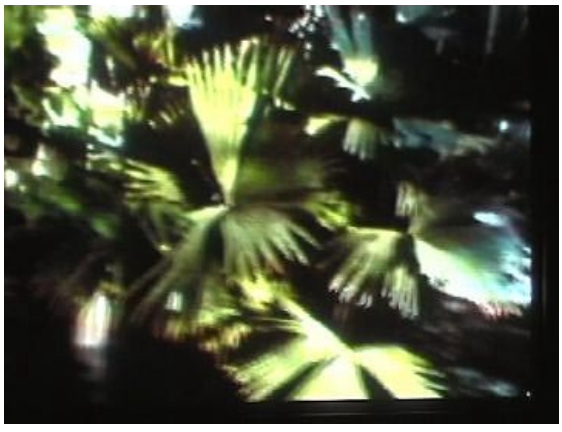


Figure 6 The BENOGO Botanical Environment

6.3 Talk-aloud analysis

The talk-aloud approach we used was part of a raft of techniques employed during data gathering in demo 1. 10 participants took part in this section of the BENOGO test while only half that amount took part in the benchmarking activities, which are an on going process. The talk-aloud method we used for this study was to a certain extent treated more like a live interview than the ones we had done previously [22]. This was largely due to the restrictive nature of the first Demo where no movement or interaction was actually possible within the virtual environment. We could not effectively observe interactions with the virtual environment; therefore this first study was focused on a description of the significant aspects of the environment highlighting technical problems with the VE but limited in exploring the concept of presence.

The videos were first of all transferred from tape to hard disk for storage and easy access. They were then viewed a number of times to promote immersion in the data and a transcript of participants comments was rendered along with notes on aspects of visual, audio and timing of events. Analysis of the talk aloud method resulted in a number of factors that consistently recurred in all of the sessions. Broadly speaking these were grouped into the three categories that the talk aloud questions were designed to enquire about.

1. Description: The descriptive level of the environment: recognisable objects and features of the environment, trees, plants, bridge etc (6.4).
2. Significance: The personal subjective engagement with the environment. Feelings of calmness, pleasantness, lack of atmosphere and humidity, memories of holidays etc (6.5).
3. Realness: The technical limitations of the environment: cables, HMD, resolution etc (6.6).

Additionally there were two other main areas of interest that arose from the talk aloud sessions that were not considered before the tests, Movement (6.7) and Sound (6.8).

6.4 Descriptions, things and objects

In the BENOGO environment, the types of elements that people could see were obviously identifiable despite the resolution problems that arose due to the technical limitations. Interestingly enough participants often identified these technical problems as things that they could see in this descriptive section as if they were objects in the environment (e.g. "I can see stereo"). Every single participant commented on a computer generated sculpture that had been added to the world and how odd it looked.

In the real botanical garden in Edinburgh similar types of description occurred where participants identified particular objects such as trees, plants, water, the building etc. As well as these a number of other things were highlighted in the descriptions of the real world. Fish, birds, signposts, heat, humidity and people were all existent in the real world but not present in the BENOGO environment. The only time any of these things was mentioned in the BENOGO environment was to point out their absence. (Note: participants are labelled r-real, b-BENOGO)

"I see a garden, with a bridge and an object, looks like coming from a leaf, staying in the middle, then I see the sun on the leaves. I hear some water. I see the roof." *Participant b2*

"There is no moisture in the air, in my breathing or sensing on my skin. That's one of the things I'm missing" *Participant b10*

6.5 Significance and Memory

In terms of personal responses to the environments it was in the real world that much more reference to significance and memory occurred. Participants were often reminded of other places that

they had been. Other botanical gardens, gardens in general or places with hot climates. Cultural references to films such as the Jungle Book, Tarzan and general jungle films were mentioned and personal memories of holidays, family members and in the case of two Greek participants, home were also mentioned. In the BENOGO environment very little of this type of data was uncovered. There were some mentions of memories of other botanical gardens and holidays but very little that was as vivid as those in the real environment. In the BENOGO environment there were a few mentions of games and gaming related comments that were not present in the real environment.

"It reminds me of Kew Gardens I went there when I was younger, the other thing is the heat and the condensation it reminds me of a shower. Its very relaxing and quiet" *Participant r1*

"It reminds me of a place, a museum in Copenhagen which has a kind of indoor garden like this. It's not the same actually but it sounds very much the same… it reminds me of being on a holiday in a different place. Actually it doesn't remind me of a rain forest although it could be but there's too much light in here." *Participant b10*

6.6 Realness

In the BENOGO environment comments about realness were almost always couched in relative terms. Most people understood or pinpointed resolution problems that made the visuals seem unreal. At the same time most of the participants said that it 'looked pretty real' particularly in relation to other types of VR.
In the real environment the same thing happened but this time in reverse. Everybody understood that they were in a real environment and that they could see real things but the man-made construction of the physical environment brought out comments such as 'fake' or 'unnatural' that seemed to impinge upon participants sense of realness.

"It is an artificially created real place. Everything around me is real I can touch it. It is tangible." *Participant r1*

"I think the way I see through the glass in here or whatever, is a bit blurry especially when I move quickly, but I think that it looks like a place that is here and I am looking through something." *Participant b14*

6.7 Movement

Participants in the real world had much more freedom to move around the environment. In the BENOGO environment attempts to move and mentions of wanting to move were quite common across most participants but these were physically restricted by the cables of the HMD, and technically by the 1 point of view REX (Region of Exploration) of the environment.

"I get the feeling of being attracted to walking over the bridge or trying to step down on some other place maybe walk round, to explore it even more. This possibility of being able to move around this place would enhance the feeling of being there." *Participant b10*

6.8 Sound

Sound also featured quite highly in both environments. In the BENOGO environment many participants commented on the sound and its suitability to the visuals. However many realised that the sound was not necessarily connected to the visuals even although there was something directional about it. Comments often arose about cars outside, birds and the noise of water in the environment. These were sometimes accompanied by comments about the water not moving visually while it sounded like it was or no movement in the trees where birds might be. In the real environment sound comments were restricted to comments on the water, the humidifier being turned off and on, and the sense of quiet in the space.

"Sound, sound is very spatial it's location based." *Participant b8*

"I can hear this bird's cry somewhere in the soundscape. So I, for a while, actually try to locate the bird. It seems to be impossible for me." *Participant b10*

6.9 Discussion

In terms of the model presented earlier this BENOGO study points towards some interesting aspects of a semiotics of Virtual Environments. Although unlike our other studies [22], due to restricted interaction, it is still possible to see relationships with aspects of the model in our data. For example in the descriptive level we get a sense of artefacts in the environment trees, plants, bridge etc. These are examples of simple denotation active in the representational images of the environment. Furthermore there are examples of a more connotative semiotics present in the environments. These are captured through focusing on elements of significance and memory. In short this is the territory of the Umwelt. Here an environment such as a botanical garden can trigger a sense-making semiosis that allows cultural references to jungle movies or memories of holidays to take place. A link then is established between encountering the phenomena of the environment and the signification process within the Umwelt.
Although we are still early on in the BENOGO project, we consider that this approach has been useful in exploring the differences between real and virtual places. What it has highlighted, in a similar way to the Photoshop studies, is the relative difference in the richness between real and virtual environments. Denotative aspects come through strongly in both but connotative aspects are much stronger in the real world. This may have something to do with embodiment in terms of extra channels of sensory input that only occur through actually being in a place. It may also have something to do with the technical limitations of the technology that interferes with this feeling presence e.g. low resolution heavy HMD, etc. As the BENOGO project develops these are areas that we are keen to explore in terms of the concept of mediation and the semiotic model presented.

7. CONCLUSION

In this paper we have presented our approach to understanding interactive systems as new media. The reason why books such as those by Dourish and McCullogh and Manovich are appearing is because human-computer interaction is changing in the light of new media. Dourish refers to tangible computing and to social computing and the changes that these are bringing to the ways we think about HCI. We are currently working with photo-realistic virtual reality in order to investigate the notion of presence [5]. How can we design systems so that people feel they are somewhere else? As part of BENOGO we are contributing to developing measures of presence and aim to provide designers with a pattern language for designing for presence. The work

presented here contributes to this notion as we seek ways of measuring the 'amount' of mediation in an interaction on various dimensions such as fidelity, interest level, concentration level and so on. This is a radical departure from previous approaches to presence that have concentrated on physiological measures.

We are also keen to explore the new information spaces that are being created through pervasive, distributed computing environments. Here Benyon has already characterised a new HCI, concerned with the navigation of information spaces [8,10], from a semiotic perspective. Looking to lessons from architectural semiotics, interior and garden design we are looking to the design of physical environments with many embedded information and communication devices. The information space is, thus, built into the environment and people are in a very real sense inside an information space. Theirs is a zone 1 medium that they will shape and form into an environment within which they can engage in activities.

The semiotic analysis of information spaces provides an alternative and, we believe, useful perspective on interaction with and through new media. Designers have this one-off chance, the 'one shot message', to communicate with their users. This message is the medium with which and through which people interact. This is part of the intersubjectivity that Dourish deals with, seeing 'communication between designer and user as medium' [14]. But the medium is made of interactive systems and we have developed a semiotic model of interactive systems that captures the temporal as well as spatial relations between signs (information artefacts), their denotations and their connotations. The individual also brings a host of background knowledge and interpretations to the interactions in the form of his or her 'Umwelt'. Here we see connections with the perception/action loop that characterises much phenomenology and with the interpretations through blends and metaphors suggested by Lakoff and Johnson [17].

The semiotic analysis lets us go beyond the denoted meanings of things and asks us to consider the connotations and cultural effects that designs have. We are increasingly living in a physical world augmented by virtual displays, and populated by interconnected information and communication devices. This new medium needs new approaches to assist designers and the semiotics of interactive systems is such an approach.

8. ACKNOWLEDGEMENTS

Susan Turner has contributed much to these ideas during discussions, as have the partners at BENOGO see http://www.benogo.dk/.

9. REFERENCES

[1] Allot, R., Language and the origin of semiosis. In W. Noth (Ed.), *Origins of Semiosis: sign Evolution in Nature and Culture*, Morton de Gruyter, Berlin, 1994.

[2] Andersen, P.B., *A Theory of Computer Semiotics*, Cambridge University Press, Cambridge, 1990.

[3] Andersen, P.B., A semiiotic approach to programming. In B.H. Peter bogh Andersen, Jens F. Jensen (Ed.), *The Computer as Medium*, Cambridge University Press, Aarhus, 1993, pp. 16-67.

[4] Andersen, P.B., What Semiotics can and cannot do for HCI, *Knowledge-Based Systems*, Volume 14 (2001) 419-424.

[5] Arnspang, J., Benyon, D. R., Fahle, M. W., Granum, E., Madsen, C. W., Pajdla, T., Peleg, S., Smyth, M., Turner, P., Turner, S. and Weinshall, D., An Investigation Into Virtual Representations Of Real Places., *In Proceedings of PRESENCE 2002:*, The 5th Annual International Workshop on Presence, Universidade Fernando Pessoa Porto, Portugal - October 9,10,11, 2002.

[6] Barbosa, S., et al, Direct and Indirect user-to-developer messages through communicability evaluation. *Representational Support for User Developer Communication workshop, INTERACT'99*, 1999.

[7] Benyon, D., Beyond the Metaphor of Navigation in Information Space, *chi2000* (2000).

[8] Benyon, D., The new HCI? Navigation of information space, *Knowledge-Based Systems*, Volume 14 (2001) 425-430.

[9] Benyon, D., & Hook, K., Navigation in Information Spaces: supporting the individual. In S. Howard, Hammond, j., Lingaard, G. (Ed.), *Human Computer Interaction: INTERACT'97*, Chapman & Hall, 1997.

[10] Benyon, D.R., Cognitive Ergonomics as Navigation in Information Space, *Ergonomics 41 (2) Feb* (1998) 153 - 156.

[11] de Souza, C.S., et al, A Semiotic Engineering Approach to User Interface Design, *Knowledge-Based Systems*, 14 (2001) 461-465.

[12] De Souza, C.S.P., R.O.; Carey, T, Missing and Declining Affordances: Are these Appropriate Concepts?, *Journal of the Brazilian Computer Society*, vol.6 (2000).

[13] Deely, J., Umwelts Semiootika osakonna kodulehekulg, *Semiotika 134*, special volume about Jakob von Uexkull (2001) 125-135.

[14] Dourish, P., *Where the Action Is*, MIT Press, Cambridge, 2001.

[15] Eco, U., *A theory of Semiotics*, Indiana University Press, Indiana, 1976.

[16] Kull, K., On semiosis, Umwelt, and semiosphere, *Semiotica*, vol. 120 (1998) 299-310.

[17] Lakoff, G.J., M, *Philosophy of the Flesh*, 1999.

[18] Manning, P., *Semiotics and Fieldwork*, Sage Publications, California, 1987.

[19] Manovich, L., *The Language of New Media*, MIT Press, 2001.

[20] McCullogh, M., *Abstracting Craft, The practiced digital hand*, MIT Press, Cambridge Massachusetts, 1996.

[21] McLuhan, M., *"Understanding Media: The extensions of Man"*, Routledge, London, 1994.

[22] O'Neill, S., Benyon, D. R. and Turner, S., The Semiotics of Interactive Systems., *To appear in Cognition, Technology and Work* (forthcoming).

[23] O'Neill, S., Benyon, D. R. and Turner, S., Semiotics and Interaction Analysis. proceedings of *ECCE 11*, Catania Sicily, 2002.

[24] Prates, R., de Souza, C, and Barbosa, S, A method for evaluating the communicability of User Interfaces, *Interactions* (2000) 31-38.

[25] Prates R, d.S.C., and Barbosa S, A case Study for Evaluating Interface Design through Communicability. *ACM Designing Interactive Systems, DIS'2000.*, Brooklyn, NY., 2000.

[26] Sebeok, T.A., *The sign and its Masters*, University of Texas Press, Austin, 1979.

[27] Winograd T and Flores, F., *Computers and Cognition.*, 1986.

The Fly's Eye: Live Spatial Analysis And Deconstruction Of The Video Image

Andrea Polli, Assoc. Professor
Hunter College, Film and Media
695 Park Ave.
New York, NY 10021
apolli@hunter.cuny.edu

ABSTRACT

The Fly's Eye project (2002) creates an animated document of both space and time and draws inspiration from the structure, function and significance of the eye of the fly and other processes of vision. In *The Fly's Eye*, the history of a public space or film is built in layers of position and image.

Keywords

Video analysis, installation, art, interactive, film.

1. ARTIST'S STATEMENT

My work has been placed in the realm of art, in the realm of science, and in the realm of cultural crossover between the two. Because I have been interested in creating interactive experiences, the science informing my work and research has primarily been in the area of perception, but all my projects are hybrids that feed into multiple areas of research. *Like The Fly's Eye* project, my work often involves the development of new physical human-computer interaction devices, and offers a new 'reading' of physical information that brings to view unfamiliar aspects of the information.

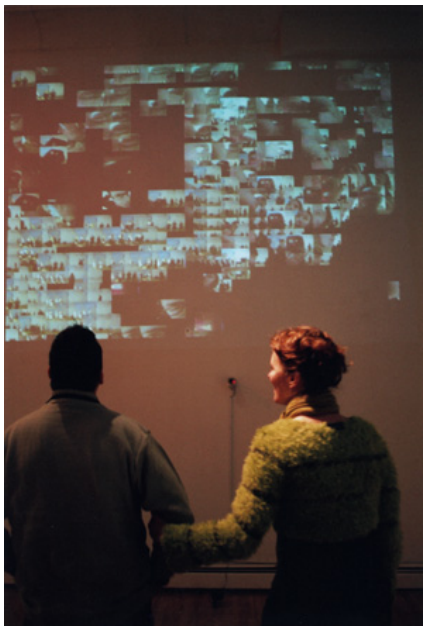


Figure 1. Visitors to Apex Art Gallery in Tribeca, new york viewing the installation January, 2003.

2. INSTALLATION

In *The Fly's Eye* installation, multiple images are projected in a public space based on the movement of viewers in the space. *The Fly's Eye* 'watches' the viewer in the space while the viewer simultaneously enjoys some control and direction of the location of the image. Each time the viewer changes position, the live video feed moves and a visible trail is left on the projected image (see figure 1).

3. TECHNICAL DESCRIPTION

The Fly's Eye consists of a computer system designed to perform a real-time spatial analysis and deconstruction of a live or pre-recorded video using a custom designed interface. Each video frame is tracked and analyzed according to the location of light, color, or motion in the frame. A copy of each video frame is placed in a grid according to the results of the analysis, and a live animation is created (see figure 2).

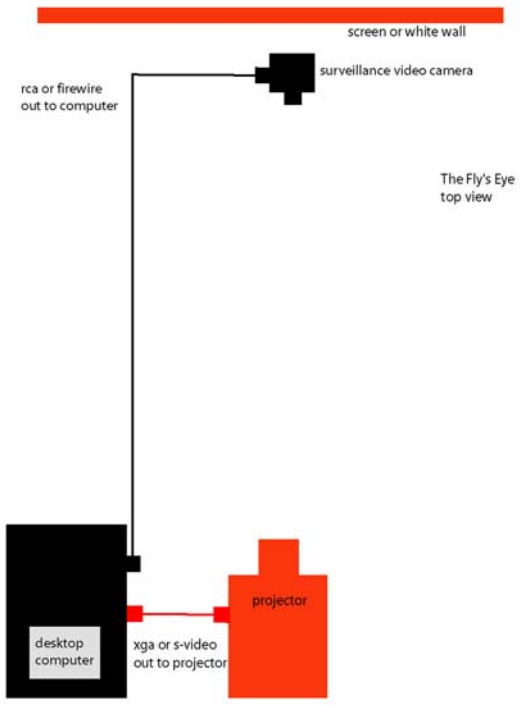


Figure 2. The Fly's Eye basic equipment configuration.

3.1 Technical Summary

- The computer performs a real-time spatial analysis of a live or pre-recorded video
- Video frames are tracked and analyzed according to the light, color, or motion in the frame.
- A copy of each frame is placed according to the results of the analysis, and a live animation is created

4. DIGITAL PRINTS

The Fly's Eye system can also be used to analyze color and lighting in films. To show the results of the analysis of several films, large format digital prints have been created.

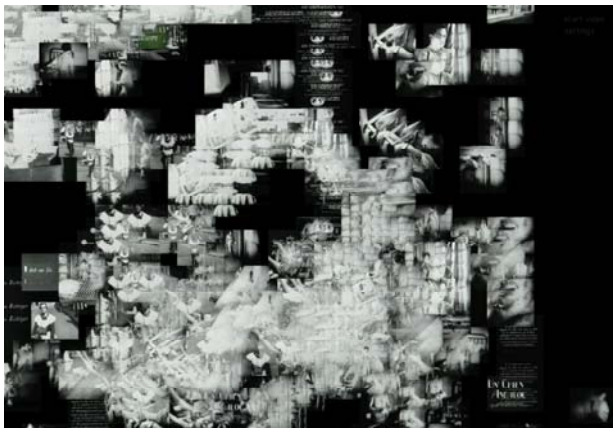


Figure 3. Detail of a lighting analysis of *Un Chien Andalou*.

These digital prints include: a lighting analysis of Fellini's *8 1/2* in which the print is divided into a grid of 28 squares, each documenting a ten minute section of the film layered over the previous ten minutes; an analysis of the location of the color red in five minute sections of the film *Moulin Rouge*; and a lighting analysis of Bunuel's *Un Chien Andalou* in which the print is divided into three rectangles each documenting a 5-10 minute section of the film (see figure 3). Prints vary in size but are approximately 66" X44"

Figure 4. An example of a movement portrait.

4.1 Movement Portraits

Another version of *The Fly's Eye* system has also been used to create a series of movement portraits. These portraits, created during a sitting of about 2 minutes, are built based on the movement of the subject. As the portrait subject moves, the portion of their image that moved is placed on the screen. Several digital prints have been created as movement portraits (see figure 4).

5. ACKNOWLEDGMENTS

The installation and/or digital prints have been shown at:

1. Apex Art Gallery in Tribeca, New York
2. Swanson-Cralle Gallery in Louisville, KY
3. Politecnico di Milano University in Milan, Italy
4. The Kunstgewerbe Museum in Berlin, Germany
5. RCA Visual in St. John's Newfoundland
6. The Aronoff Center in Cincinnati, OH
7. SIGGRAPH 03 in San Diego, CA

Thanks to: Stuart Ewen, Birgit Rathsmann, and Chuck Varga for assistance with the installation, digital prints and documentation.

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

A Game Pidgin Language For Speech Recognition In Computer Games

Tarashankar Rudra
Charles Sturt University
Bathurst, NSW
Australia -2795
trudra@csu.edu.au

Manolya Kavakli
Charles Sturt University
Bathurst, NSW
Australia -2795
mkavakli@csu.edu.au

Terry Bossomaier
Charles Sturt University
Bathurst, NSW
Australia -2795
tbossomaier@csu.edu.au

ABSTRACT
Today very little progress has been made in the field of Computer Science towards the development of robust Natural Language Processors. Developing a Computer Pidgin Language (CPL) with limited vocabulary and simple set of grammatical rules can be an effective approach to tackle the problem of humans interacting with computers. A CPL [7] is a new spoken language, which is taught to the user and is efficient for dialogues with the computer. In this paper an attempt has been made to develop a Game Pidgin Language (GPL) with limited vocabulary, grammar and syllables for use with speech interactive Computer Games. The GPL is illustrated with a set of vocabulary that is designed to optimise the efficiency of automatic speech recognizer (ASR).

Keywords
Pidgin languages, Information theory, Speech Recognition, Extensible Markup Language.

1. INTRODUCTION
The earliest record of the use of "pidgin" languages dates back to the Middle Ages when the European crusaders and traders of the eastern end of Mediterranean used it. Due to the dominance of French among the crusaders the language was known as *Lingua Franca*, which denotes any language that is used as a medium of communication among people having no other language in common. When a speech community endorses a pidgin language, it becomes a *Creole*. For a language to be pidgin, it must satisfy two conditions [5] –

1) It must have a sharply reduced grammatical structure and vocabulary.
2) The language must be native to none who use it.

The idea for the simplification of natural language has been envisaged for the communication of complex concepts with the help of simple expressions. Currently controlled language applications [e.g., 4, 13] have been developed for computational linguistic. Most controlled languages are for people whose profession is to write and who can be trained in the use of this new language. We attempt to design a language for speech interactive computer games. We use a simplified language because we want to minimize repeated words.

"CPL or Computer Pidgin Language is a radical departure from the normal approach to Speech Recognition Systems. CPL is inspired by a frustration at a perceived lack of progress in Spoken Language Research over the last 20-30 years" [7]. Hinde & Belrose believe that systems that only understand people 85% of the time are hardly usable, so speech recognition is very much a last resort technology or a curiosity. "One of the inspirations for thinking about CPL as an approach to spoken language recognition is observing the evolution of handwriting recognition." [7].

Computer games use 'barks', which is slang developed for the communications between game agents. The GPL we have developed will enhance the group bonding especially in multi-player games.

Speech recognition, or speech-to-text conversion involves capturing and digitising the sound waves, converting them into basic language units or phonemes, constructing words from phonemes, and finally contextually analysing the words to ensure correct spelling for words that sound alike (such as site and sight). The reverse process takes place in Speech synthesis or text-to-speech conversion.

Due to the absence of systematic cues marking word boundaries in continuous speech [2], we will incorporate short but maximum meaning bearing words with minimal syllables and grammatical constraints in the new language. Since Computer Games are primarily a source of fun and entertainment, our attempt is to create funny sounding but quickly memorisable words.

In this paper, we use Information Theory [15] to analyse four aboriginal languages and a sample pidgin language that we have developed.

Our hypothesis in this study is to show that the difference in entropy and perfect information content is minimum and this characteristic makes it rich in vocabulary and simple in grammar.

Information Theory, a branch of probability has two primary goals [3]–

1) Development of fundamental theoretical limits on the achievable performance when communicating a given information source over a given communication channel using coding schemes from within a prescribed class.
2) Development of coding schemes that is reasonably good in comparison with the optimal performance given by the theory.

Entropy provides the information of a random process about itself and measures the information content or uncertainty of 'x'.

Entropy (H(X)) is given by –

Formula - I: $\mathbf{H(X) = - \Sigma_{x \epsilon Ax} P(x)*Log_2(P(x))}$ [3].

Here an ensemble 'X' is a random variable 'x' with a set of possible outcomes, $Ax = \{a_1, a_2, \ldots\ldots a_i\}$, having probabilities $Px = \{P_1, P_2, \ldots\ldots P_i\}$, with $P(x= a_i) = P_i$, where $P_i > 0$ and $\Sigma_{x \epsilon Ax} (Px) = 1$.

In our study, *Ax* is a set of distinct words appearing in the paragraph, the set P*x* consists of probability of occurrence of distinct words in the paragraph.

Perfect Information content ($H_0(X)$) is a lower bound for the number of binary questions that are guaranteed to identify the outcome. It is given by –

Formula - II: $\mathbf{H_0(X) = Log_2|\, Ax|}$ [12].

In this paper we have focused on developing a framework of a Computer Pidgin Language (CPL) illustrated with a set of vocabulary for speech interactive computer games. The emphasis of the vocabulary is on having minimum difference between ***Entropy*** and ***Perfect Information*** content of the language. Our endeavor is also to have much lower value of Entropy of GPL than English. We have developed an Extensible Markup Language (XML) called GPLXML to structure our grammar and demonstrated its use with an instance. An XML defines a set of rules, which identifies how we can define tags that separate a document into individual parts and subparts [17].

2. INFORMATION CONTENT IN THE USE OF ABORIGINAL LANGUAGES

There were approximately 200 aboriginal languages in Australia, out of which nearly 100 are in use [1]. The characteristics of Aboriginal languages are [8] -

1) They are rich in vocabulary.
2) Complex words can be formed by compounding, reduplication or by using suffixes.
3) They normally do not have sounds of f, v, s, z and sh.
4) They have terms for every species of animal and plant in their environment.
5) They have elaborate vocabulary in the area of kinship.
6) They have complex syntax and word building processes.
7) They tend to have similar sets of speech sound and share numerous grammatical features but differ greatly in vocabulary.
8) All of them usually have the sounds of p, b, t, d, k and g.

In this paper, our purpose is to focus on information content related to semiotics rather than phonemes. The first four characteristics of Aboriginal languages prompt us to incorporate some of their vocabulary in our GPL along with others inherited from different languages of the world.

Reducing the number of phonemes has considerable advantages for speech recognition. Recent work suggests that languages such as Italian, which are largely phonetic and have fewer phonemes than English, have a much lower rate of dyslexia [6]. For our purposes we can make recognition faster and more robust. However, the focus of this paper is not on the phonetic representation as such, but the semiotic qualities of the pidgin language. There are many ways of calculating the entropy of a block of text. Words have strong serial correlations, which affect the joint entropy. In this first study we have elected to look at word frequencies only.

An example showing the calculation procedure for extracting $H(X)$, $H_0(X)$ parameters are given below, the data is listed in Table 3 –

a) Language 1: English

b) Test paragraph: The Guugu Yimithirr people lived in Northern Queensland, in an area northwest of Cooktown, where Captain Cook first landed in Australia and first met an aboriginal language. So, the first aboriginal language they heard was Guugu Yimithirr. The remaining Guugu Yimithirr live now in Hopevale, and there around 100 speakers of the language left. Names correspond to traditional areas where the features named by colonialists appear in maps.

c) Total number of words: 68

d) Number of unique words: 47

e) Occurrence of distinct words and their probabilities in b) are listed in table-1 -

Table-1

Word	Occurrence	P_i
The	5	0.0735
Guugu	3	0.0441
Yimithirr	3	0.0441
people	1	0.0147
lived	1	0.0147
in	5	0.0735
Northern	1	0.0147
Queensland	1	0.0147
an	2	0.0294
area	1	0.0147
Northwest	1	0.0147
Of	2	0.0294
Cooktown	1	0.0147
where	2	0.0294
Captain	1	0.0147
Cook	1	0.0147
first	3	0.0441
landed	1	0.0147
Australia	1	0.0147
and	2	0.0294
met	1	0.0147
Aboriginal	2	0.0294
language	3	0.0441
So	1	0.0147
they	1	0.0147
heard	1	0.0147
was	1	0.0147
Remaining	1	0.0147

live	1	0.0147
now	1	0.0147
Hopevale	1	0.0147
there	1	0.0147
Around	1	0.0147
100	1	0.0147
speakers	1	0.0147
Left	1	0.0147
Names	1	0.0147
Correspond	1	0.0147
To	1	0.0147
Traditiona	1	0.0147
areas	1	0.0147
features	1	0.0147
named	1	0.0147
By	1	0.0147
colonialists	1	0.0147
appear	1	0.0147
maps	1	0.0147

f) Entropy: 5.32, by applying formula I.

g) Perfect Information: 5.55, by applying formula II.

a) Language 2: Guugu Yimithirr

b) Test paragraph: Guugu Yimithirr herria bizi zen Queensland iparraldean, Australian, Cook Kapitainak lur hartan oina lehenbizikoz jarri zuen tokian. Beraz, hizkuntza hau izan zen Australian zuriek entzun zuten legena. Oraingo Guugu Yimithirr jendea Hopevale deritzon tokian bizi dira, eta hiztunak 100 inguru izan litezke. Gure zerrendako izenak dira eremu tradizionalen izenak, non kolonialistek izendatutako tokiak ageri diren mapetan.

c) Total number of words: 56

d) Number of unique words (UW): 47

e) Occurrence of distinct words and their probabilities in b) are listed in table-2 -

Table-2

Word	**Occurrence**	**P_i**
Guugu	2	0.0357
Yimithirr	2	0.0357
herria	1	0.0179
bizi	2	0.0357
zen	2	0.0357
Queensland	1	0.0179
Iparraldean	1	0.0179
Australian	2	0.0357
Cook	1	0.0179
Kapitainak	1	0.0179
lur	1	0.0179
hartan	1	0.0179
oina	1	0.0179
Lehenbizikoz	1	0.0179
jarri	1	0.0179
zuen	1	0.0179
tokian	2	0.0357
Beraz	1	0.0179
Hizkuntza	1	0.0179
hau	1	0.0179
izan	2	0.0357
zuriek	1	0.0179
entzun	1	0.0179
zuten	1	0.0179
legena	1	0.0179
Oraingo	1	0.0179
jendea	1	0.0179
Hopevale	1	0.0179
deritzon	1	0.0179
dira	2	0.0357
eta	1	0.0179
hiztunak	1	0.0179
100	1	0.0179
Inguru	1	0.0179
litezke	1	0.0179
Gure	1	0.0179
Zerrendako	1	0.0179
izenak	2	0.0357
eremu	1	0.0179
Tradizionalen	1	0.0179
non	1	0.0179
Kolonialistek	1	0.0179
Izendatutako	1	0.0179
tokiak	1	0.0179
ageri	1	0.0179
diren	1	0.0179
mapetan	1	0.0179

f) Entropy: 5.49, by applying formula I.

g) Perfect Information: 5.55, by applying formula II.

Analyzing the data of Table 3 to 6, we find that our GPL (table 8) is most suitable as a new pidgin language for games and Guugu Yimithirr is least suitable.

The following tables provide the statistical measures of information in Aboriginal languages [1] based on the principles of Information theory.

In the following tables, UW = unique words; H = entropy; H_0 = perfect information.

Table-3

English				Guugu Yimithirr			
Words	**UW**	**H**	**H_0**	**Words**	**UW**	**H**	**H_0**
68	47	5.32	5.55	56	47	5.49	5.55

Table-4

English				Arabana-Wangkangurru			
Words	**UW**	**H**	**H_0**	**Words**	**UW**	**H**	**H_0**
57	47	5.40	5.55	36	33	5.00	5.04

Table-5

English				Dyirbal			
Words	**UW**	**H**	**H_0**	**Words**	**UW**	**H**	**H_0**
93	68	5.84	6.09	64	57	5.76	5.83

Table-6

English				Yagara / Yugambeh			
Words	**UW**	**H**	**H_0**	**Words**	**UW**	**H**	**H_0**
59	52	5.63	5.70	49	41	5.26	5.36

3. MODEL OF A PIDGIN LANGUAGE

3.1 GPL Vocabulary

Advantages of GPL in computer games are –

1) The more limited vocabulary and simple grammar makes speech processing across many ethnic backgrounds much easier.
2) Games move fast, and short simple utterances are more appropriate.
3) Peer groups and sub-cultures love to have their own "cool" vocabulary. An extension to the work we discuss will look at adaptive mechanisms for modifying the language by the game players themselves.
4) It has great significance in building the cognitive models for the animats in the game. The mini-language defines the cognitive framework within which they operate and determines also the sophistication (and feasibility) of their world view.

The words in the sample GPL dictionary has been compiled from words in English, Australian Aboriginal languages and some International languages, they are –
Arabic, English, Bengali, Gunyah [14], Hindi, Kamilaroi [11], Portuguese, Russian, Slang English, Wiradjuri [9].

The sample GPL dictionary has the following words –

Here N = noun; Pron = pronoun; V = verb; Adj = adjective.

Table-7

Word	Meaning	Type	Phrase	Source
Ma	I, my, me, myself	N		
Gaba	Man	N		Wiradjuri
Tum	You, your	pron		Hindi
Bingo	Expression of joy	V	Y	English
Fat	Expression of despair	V	Y	
Aka	Master	N		Arabic
Kid	an inferior	N		slang
Goofy	stupid	Adj		slang
Yea	Yes	N		English
Nay	No	N		English
Wa	What, when, where, which, how	V		
Wana	do you want?	V	Y	
Wata	What is it?	V	Y	
Wamba	Who is it/he?	V	Y	Kamilaroi
Wara	Where is/are it/you	V	Y	
Waay	look out	V	Y	Wiradjuri
Birra	move/went away	V		Wiradjuri
Gaja	get away	V	Y	Wiradjuri
Gaa	to take	V	Y	Kamilaroi
Eta	it/he/she is a	Pron		Russian
Bom(b)	Bomb	N		
Buma	to hit, kill	V	Y	Kamilaroi
Samba	dance	N		Portuguese
Jet	Aircraft	N		
Noka	Ship	N		Bengali
Villi	Enemy	N		
Buddy	Friend	N		English
Limbo	trouble	N		slang
Gali	Water	N		Kamilaroi
Kola	drink	N		
Rivi	River	N		
Croc	Crocodile	N		English
Duma	House	N		Russian
Humpy	shelter	N		Gunyah
Dud	dumb, slack	Adj		English
Gubi	swim	V	Y	Kamilaroi
Guju	jump	V	Y	
Gumo	Climb	V	Y	
Zoom	to fly, drive, run, fast movement	V	Y	English
Yami	Food	N		
Yaki	dirty: adj	Adj		
Wee	sit	V	Y	Wiradjuri
Kam	Work	N		
Jumbo	huge	Adj		English
Dagi	to pierce with	V	Y	

	sharp object	
i (é)	Is	V

The dictionary we have developed is an example that illustrates the idea of developing a Pidgin language for Game play.

3.2 GPL Grammar

The GPL has a very limited set of grammatical rules. It primarily has eight rules –

1) The Sentence may begin with a noun or noun phrase followed by Verb.
2) A Sentence may begin with a Verb Phrase followed by a Noun.
3) An Adjective is a valid sentence.
4) A Noun is a valid sentence.
5) A Verb Phrase is a Valid Sentence.
6) The auxiliary 'i' (é) is added as suffix to a noun that does not end with 'i', 'a' or 'y' sounding alphabet.
7) There is no tense in this language as Computer games are played in real time.
8) The language has no gender.

A valid sentence is any word or phrase that satisfies conditions 1 to 5 and is composed of word or words within the GPL dictionary. A valid sentence may generate a response from the system.

Examples of some GPL sentences:

1) I went home. – Ma birra doma.
2) I had food and water. – Ma yami gali gaa.
3) Shoot the Aircraft. – Jet buma.
4) Jump in the river. – Rivi guju.
5) He is a dumb guy. – Eta dud.
6) A crocodile is swimming in the dirty river. – Kroki gubi yaki rivi.
7) Look out for the enemy. – Waay villi.
8) Gosh! You killed a friend. – Fat! tu buddy buma.
9) Do you want a coke? – Wana kola?
10) Where are the enemy aircrafts? - Wara villi Jet?
11) Run for shelter! – Zoomi humpy!

3.3 Document Type Definition (DTD) of the GPL Grammar

A DTD defines rules for validating an XML document using Backus-Naur-Form grammar to identify, which elements are valid for a particular XML document and which attributes are then valid to be used with each of those elements [17]. The popularity of DTD is due to its ease of development coupled by the availability of more Software's for validating and testing the conformance of XML documents.

The GPLXML we have developed will be used for word recognition. The simple tags will make it easier to synthesize utterances from cognitive models. It has the following DTD -

```
<?xml version="1.0" standalone="yes"?>
<!ELEMENT GPLXML (grammar)>
<!ELEMENT grammar (nouns, verbs, adjectives, pronouns)>
<!ELEMENT nouns (noun)+>
<!ELEMENT noun (word, meaning, feature+)>
<!ELEMENT verbs (verb)+>
<!ELEMENT verb (word, meaning, feature+)>
<!ELEMENT adjectives (adjective)+>
<!ELEMENT adjective (word, meaning, feature+)>
<!ELEMENT pronouns (pronoun)+>
<!ELEMENT pronoun (word, meaning, feature+)>
<!ELEMENT word (#PCDATA)>
<!ELEMENT meaning (#PCDATA)>
<!ELEMENT feature (phoneme, emotion?)>
<!ELEMENT phoneme (#PCDATA)>
<!ELEMENT emotion (#PCDATA)>
<!ATTLIST noun type CDATA #REQUIRED>
<!ATTLIST verb phrase (yes|no) "yes">
<!--"prosody" features as attributes in "emotion", the
      prob attribute provides accuracy in judgement -->
<!ATTLIST emotion pitch CDATA #IMPLIED range CDATA #IMPLIED
prob CDATA #IMPLIED>
```

The GPL DTD has `grammar` as its root node. The `grammar` has `words` along with their `meanings` and `features` classified in `nouns`, `verbs`, `adjectives` and `pronouns`. The `feature` has `phoneme` and `emotion` tags that store `phonetic` representation and `emotion` features of every `word`. The `emotion` tag has the prosody features of `pitch` and `range` inherited from Java™ Speech API Markup Language [16] as attribute along with `prob` that stores the probability of accuracy of judgement of emotion. The `Prob` attribute will provide flexibility of programming to the Game developer.

XML schema definition language (XSD) is a powerful but flexible document definition language that provides control not only over elements and attribute existence, content and order but also specifies when and how elements and attributes can be used along with the content of attribute based on the position of attribute elements within the document hierarchy [17]. The work on XML Schema began in 1999 and got recommendation status in May 2001 [10].

The GPL Schema for the GPL grammar is as follows:

<? xml version="1.0" encoding="utf-8"?>

<xsd:schema xmlns:xsd=http://www.w3.org/2001/XMLSchema>

<xsd:element name="GPLXML">

```
<xsd:attributeGroup name=”prosody”>
  <xsd:attribute name=”pitch” type=”xsd:string”
      use=”optional”/>
  <xsd:attribute name=”range” type=”xsd:string”
      use=”optional”/>
  <xsd:attribute name=”prob” use=”optional”>
    <xsd:simpleType >
      <xsd:restriction base=”xsd: nonNegativeInteger”>
        <xsd:minInclusive value=”0”/>
        <xsd:maxInclusive value=”100”/>
      </xsd:restriction>
    </xsd:simpleType >
  </xsd:attribute >
</xsd:attributeGroup>
<xsd:complexType>
  <xsd:sequence>
    <xsd:element name=”noun” minOccurs=”1”
       maxOccurs=”unbounded”>
      <xsd:complexType>
        <xsd:sequence>
          <xsd:element name=”word” type=”xsd:string”/>
          <xsd:element name=”meaning”
              type=”xsd:string”/>
          <xsd:element name=”features” minOccurs=”1”
              maxOccurs=”unbounded” >
            <xsd:complexType>
              <xsd:sequence>
                <xsd:element name=”phoneme”
                    type=”xsd:string”/>
                <xsd:element name=”emotion”
                type=”xsd:string” minOccurs=”1”
                    maxOccurs=”1” >
                  <xsd:complexType>
                    <xsd:attributeGroup ref=”prosody”/>
                  </xsd:complexType>
                </xsd:element>
              </xsd:sequence>
            </xsd:complexType>
          </xsd:element>
        </xsd:sequence>
        <xsd:attribute name=”type” type=”xsd:string”
          use=”required”/>
      </xsd:complexType>
    </xsd:element>
    <xsd:element name=”verb” minOccurs=”1”
      maxOccurs=”unbounded”>
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name=”word” type=”xsd:string”/>
        <xsd:element name=”meaning”
            type=”xsd:string”/>
        <xsd:element name=”features” minOccurs=”1”
            maxOccurs=”unbounded” >
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name=”phoneme”
                  type=”xsd:string”/>
              <xsd:element name=”emotion”
              type=”xsd:string” minOccurs=”1”
                  maxOccurs=”1” >
                <xsd:complexType>
                  <xsd:attributeGroup ref=”prosody”/>
                </xsd:complexType>
              </xsd:element>
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
      <xsd:attribute name=”phrase” use=”required“
          default=”yes”>
        <xsd:simpleType >
          <xsd:restriction base=”xsd:string”>
            <xsd:enumeration value=”yes”/>
            <xsd:enumeration value=”no”/>
          </xsd:restriction>
        </xsd:simpleType >
      </xsd:attribute>
  </xsd:complexType>
  </xsd:element>
  <xsd:element name=”adjective” minOccurs=”1”
    maxOccurs=”unbounded”>
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name=”word” type=”xsd:string”/>
        <xsd:element name=”meaning”
            type=”xsd:string”/>
        <xsd:element name=”features” minOccurs=”1”
            maxOccurs=”unbounded” >
          <xsd:complexType>
            <xsd:sequence>
```

```
                    <xsd:element name=”phoneme”
                            type=”xsd:string”/>
                    <xsd:element name=”emotion”
                    type=”xsd:string” minOccurs=”1”
                            maxOccurs=”1” >
                      <xsd:complexType>
                        <xsd:attributeGroup ref=”prosody”/>
                      </xsd:complexType>
                    </xsd:element>
                  </xsd:sequence>
                </xsd:complexType>
              </xsd:element>
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name=”pronoun” minOccurs=”1”
              maxOccurs=”unbounded”>
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name=”word” type=”xsd:string”/>
              <xsd:element name=”meaning”
                    type=”xsd:string”/>
              <xsd:element name=”features” minOccurs=”1”
                    maxOccurs=”unbounded” >
                <xsd:complexType>
                  <xsd:sequence>
                    <xsd:element name=”phoneme”
                            type=”xsd:string”/>
                    <xsd:element name=”emotion”
                            type=”xsd:string” minOccurs=”1”
                            maxOccurs=”1” >
                      <xsd:complexType>
                        <xsd:attributeGroup ref=”prosody”/>
                      </xsd:complexType>
                    </xsd:element>
                  </xsd:sequence>
                </xsd:complexType>
              </xsd:element>
            </xsd:sequence>

          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

3.4 An Instance of GPLXML

The instance of GPLXML begins with the *grammar* tag, which has *nouns*, *verbs*, *adjectives* and *pronouns* as sub-nodes. The *nouns* tag has multiple *noun*(s) within it and each *noun* has *type* as attribute along with *word*, *meaning* and *feature* as sub-nodes. There can be more than one *feature* tag per *noun* that stores phonetic representation of the *word* along with its *emotional* content as its sub-nodes. The *emotion* tag has optional attributes of *pitch*, *range* and *prob* to store the *pitch*, *frequency range* and *probability* (of judgement in the estimation of emotional content) respectively. The *verbs* tag has multiple *verb*(s) within it and each *verb* has *phrase* as attribute along with *word*, *meaning* and *feature* as sub-nodes. The rest is similar to that of *nouns* tag. The *adjectives* tag has multiple *adjective*(s) within it and each *adjective* has *word*, *meaning* and *feature* as sub-nodes. The rest is similar to that of *nouns* tag. The *pronouns* tag has multiple *pronoun*(s) within it and each *pronoun* has *word*, *meaning* and *feature* as sub-nodes. The rest is similar to that of *nouns* tag.
An instance that conforms to the GPLXML DTD is given below –

```
<!DOCTYPE GPLXML SYSTEM "GPLXML.dtd">
<GPLXML>
  <grammar>
    <nouns>
      <noun type="proper">
        <word>duma</word>
        <meaning>home</meaning>
        <feature>
          <phoneme> D UW M AH </phoneme>
          <emotion>
                joy
          </emotion>
        </feature>
      </noun>
      <noun type="common">
        <word> buddy </word>
        <meaning> friend </meaning>
        <feature>
          <phoneme> B AH D IY </phoneme>
          <emotion>
               joy
          </emotion>
        </feature>
      </noun>
    </nouns>
    <verbs>
      <verb phrase="yes">
        <word>wata</word>
        <meaning>what is it</meaning>
        <feature>
          <phoneme>HH W AH T . IH Z . IH T
          </phoneme>
```

```
        <emotion prob="90">
             Anger
        </emotion>
      </feature>
    </verb>
    <verb phrase="no">
      <word>i</word>
      <meaning>is, was</meaning>
      <feature>
        <phoneme>IY</phoneme>
      </feature >
    </verb>
  </verbs>
  <adjectives>
    <adjective>
      <word>jumbo</word>
      <meaning>huge</meaning>
      <feature>
        <phoneme>JH AH M B OW</phoneme>
        <emotion prob="80">
             astonishment
        </emotion>
      </feature>
    </adjective>
  </adjectives>
  <pronouns>
    <pronoun>
      <word>tu</word>
      <meaning>you</meaning>
      <feature>
        <phoneme>T UW</phoneme>
        <emotion>
             anger
        </emotion>
      </feature >
    </pronoun>
  </pronouns>
 </grammar>
</ GPLXML>
```

3.5 INFORMATION CONTENT OF GPL

Since the use of GPL is in short bursts, we have analyzed the examples of section 3.2 to extract the value of Entropy and perfect information of English and GPL.

Favorable values of the measures of information theory for GPL are as follows –

1. Lower value of perfect information and hence unique words in GPL over its English equivalent.
2. Lower value of entropy in GPL over its English equivalent.
3. Lower difference between the values of perfect information and entropy of GPL.

The values are listed in Table 8. From the table we can see that only 30 words of GPL can be used to convey the same information expressed by 52 words in English, resulting in lower value of Entropy of GPL over English. This signifies substantial reduction in grammar in GPL over English. We also find that the ratio of perfect information of GPL to English is less than 1 and the unique words (UW) to Words ratio is much higher in GPL than English, implying that the core of GPL lies in its vocabulary, which is one of the desired criteria.

Table-8

English				**GPL**			
Words	**UW**	**H**	**H_0**	**Words**	**UW**	**H**	**H_0**
52	38	5.05	5.25	30	27	4.71	4.75

The requirements of the game are for rather short utterances, where we would like to have a significant semiotic weight carried by every word. Thus the entropy should approach the maximum value where every word contributes to an utterance with equal probability. We show some examples of typical utterances and calculate the corresponding entropy. Note that we do not use all of the words of our vocabulary, but use this to demonstrate the procedure by which we measure word independence.

4. CONCLUSION

1) Both AL's and have low values of $H_0(X)$ and H(X) and hence number of unique words as compared to English for expressing the same information, thus we can express more information with minimum number of words by both AL and GPL.
2) Low differences in the values of $H_0(X)$ and H(X) in AL's and GPL shows that both AL's and GPL are rich in vocabulary.
3) The extremely low values and difference between $H_0(X)$ and H(X) of GPL over English signifies minimal grammar and richness in vocabulary of GPL.

We have shown in this paper how a CPL with small vocabulary with cues from aboriginal and other languages can be used to develop a GPL. Although much work is needed to be done in generating a cross continental, socio-culturally acceptable vocabulary, we believe that GPLXML will be sufficient for representing the grammar of GPL. The limitation of analysis of a language with Information theory is that it is silent about the characteristics and complexity of vocabulary in that language, hence we have focused on developing a limited set of vocabulary with least number of syllables that is bound by simple and non-rigid grammatical rules for use with speech interactive Computer Games.

5. ACKNOWLEDGMENTS

This research project is funded by Australian Research Council (ARC) through a linkage grant numbered LP0216837.

6. REFERENCES

[1] Australia Aboriginal Languages. http://www.geocities.com/Athens/9479/guugu.html.

[2] Bagou, O., Fougeron,C., Frauenfelder, U.,H., Contribution of Prosody to the Segmentation and storage of "Words" in the Acquisition of a New Mini-Language, Speech Prosody 2002 an International Conference, France, 11-13 April 2002.

[3] Gray, R., M., Entropy and Information Theory, Springer Verlag, 1990.

[4] Grover, C., Holt, A., Klein, E., Moens, M., Designing a controlled language for interactive model checking, CLAW 2000, 3rd International Controlled Language Applications Workshop,Seattle, Washington, USA, 29-30 April 2000

[5] Hall, R., A., Pidgin and Creole Languages, 1966, Cornell University Press, London.

[6] Helmuth, L., Dyslexia: same brain different languages, Science, Vol. 291, pp 2064-5, 2001.

[7] Hinde, S., Belrose,G., Computer Pidgin Language: A new Language to talk to your Computer?. http://www.hpl.hp.com/techreports/2001/HPL-2001-182.pdf.

[8] Horton, D., The Encyclopaedia of Aboriginal Australia, Vol 1,1994, Australian Institute of Aboriginal and Torres Strait Islander Studies, Australia.

[9] Hosking, D., McNicol, S., Wiradjuri, Panther Publishing and Printing, Canberra, 1993

[10] Hunter, D., Cagle, K., Dix, C., Kovack, R., Pinnock, J., Rafter, J., Beginning XML 2nd Edition, 2001, Wrox Press Ltd, UK.

[11] Kamilaroi/Gamilaraay Dictionary. http://coombs.anu.edu.au/WWWVLPages/AborigPages/LANG/GAMDICT/GAMDICTF.HTM.

[12] MacKay, D., J., C., A Short Course in Information Theory: Lecture notes 1 & 2. http://www.inference.phy.cam.ac.uk/mackay/info-theory/course.html.

[13] Pulman, S., Controlled language for knowledge representation, CLAW96: Proceedings of the first International workshop on Controlled language Applications, Belgium, March 1996, pp 233-242.

[14] Robinson, J., Voices of Queensland, 2001, Oxford University Press, Australia.

[15] Shannon, C., E., A Mathematical Theory of Communication, The Bell System Technical Journal, vol 27, pp 379-423, 1948.

[16] Sun Microsystems Inc, Java™ Speech API Markup Language Specification version 0.6, 2001. http://java.sun.com/products/java-media/speech/forDevelopers/JSML/.

[17] Williamson, H., XML: The Complete Reference, 2001, Tata McGraw-Hill Publishing Company Ltd, New Delhi.

Symbolic Activities In Virtual Spaces

Leonie Schäfer
Bozana Bokan
Uta Pankoke-Babatz
Wolfgang Prinz
Fraunhofer FIT
Schloß Birlinghoven
53754 Sankt Augustin, Germany
{name.surname}@fit.fraunhofer.de

Ava Fatah gen. Schieck
VR Centre for the Built Environment
The Bartlett,
University College London
1-19 Torrington Place,
London W1CE 6BT, UK
ucftajf@ucl.ac.uk

Amanda Oldroyd
Jon Sutton
BT Exact
Adastral Park
Martlesham, Ipswich IP5 3RE, UK
{name.surname}@bt.com

ABSTRACT
This paper presents an approach to combine concepts of symbolic acting and virtual storytelling with the support of cooperative processes. We will motivate why symbolic languages are relevant in the social context of awareness applications. We will describe different symbolic presentations and illustrate their application in three different prototypes.

Keywords
Symbolic Acting Storytelling, Collaborative work, Virtual Environments, Awareness, CSCW

1. INTRODUCTION
For effective collaborative working it is vital for teams to be able to access records of decisions made, minutes of meetings and document histories. It is also vital that new members of teams are able to catch up with what has happened in order to get a clear picture of the state of a project. Whilst many systems are available for recording changes and amendments to documents, and minutes are written recording decisions and actions at meetings, the information gleaned from these sources can be very sketchy. It can also be very difficult for team members to fully understand the context in which decisions were made or documents changed. For full understanding of what has happened it is necessary to perceive the course of events including the activities of actors [25],[29].

This paper describes three approaches on DocuDrama, three applications which offer generation of interactive narratives that are based on awareness information recorded on activities in a collaborative virtual environment. DocuDrama [26] has been developed as part of TOWER [24] a Theatre of Work Enabling Relationships, which allows project members to bcomee aware of project relevant activities as well as to establish and maintain the social relationships that intensify team coherence.

2. Challenges
In asynchronous cooperation, co-orientation needs technical support of awareness. In particular awareness [13] about the ongoing work processes and of the activities of the partners are requested similar to workspace awareness in real-time groupware [17]. Such support requires the recording of events [23] and a presentation of the episodes of action. The challenge of DocuDrama is to compile episodes from the stream of events recorded and to present them to the user in an intuitive way.

DocuDrama as a feature of the Theatre of Work focuses on the recording and replay of events. The creation of stories in DocuDrama sets up on the recording of cooperative activities in a team's shared environment. The recording of events results in history files, which contain the event information as abstract data sets. The challenges for DocuDrama in generating narratives of project histories are threefold. The first challenge is to sort and select meaningful events, to combine and aggregate this data and finally, to derive a meaning from the event sequences. The second challenge is to find a meaningful way of presenting the event data to the user. The third challenge is to present the project's history and progress in an entertaining way which captures the users' attention and conveys complex information fast and effectively.

In the following we will take a closer look on the methods of storytelling, introduce the concept of symbolic acting and describe the symbolic languages applied in the different DocuDramas.

2.1 Narratives
Narratives consist of story and discourse [10]. Story is thereby the content of narrative (what is told), while discourse is the medium that conveys the story (how is it told). A story is a sequence of events - actions and happenings - that are causally, temporally and spatially connected to each other. A story also contains characters and settings.

Stories can be communicated by many means, e.g. by language (both, oral and written), in images (both, fixed and moving), in gesture and movement.) They are present in many forms and therefore many discourses, e.g. conversation, novel, painting, film, pantomime, theatre, etc.

In DocuDrama the story is based on a sequence of events which result from activities of team members in a shared workspace. The different DocuDramas represent different discourses, all with another focus on the presentation of the story.

2.2 Storylines
The stories presented in the DocuDramas are dynamically generated. Stories told by traditional media usually follow a storyline. In film or theatre for example, there is a clear sequence of events which creates suspense and involves the user in the story[1]. Stories in DocuDrama based on a sequence of user activities in a workspace do not contain an inherent storyline. The dynamic generation of narratives which captures the users' attention and conveys complex information fast and effectively represents one of the major challenges to DocuDrama. There have

been several approaches on the dynamic generation of narratives, but no approach is known which uses a DocuDrama combination of research areas.

Temporal Links [16] introduces the idea of a flexible mechanism for replaying past or recent recordings of virtual environments within other virtual environments. Temporal Links is concerned with time, spatial and presentational relationships between the environment and the recording. Where Temporal Links focuses on replaying the past and its implications with the current environment, DocuDrama is concerned with selection and aggregation of history events and their replay depending on the user's situation.

Brooks has investigated with Agent Stories [6; 7] a model for the computational generation of narratives. This model splits the task into: defining an abstract narrative structure, collecting material and defining a navigational strategy. While Brooks offers a story design and presentation environment for non-linear, multiple-point-of-view cinematic stories, DocuDrama focuses on the automated generation of narratives by selection and aggregation of events.

2.3 Symbolic Acting

The idea of Symbolic Acting is to 'Let the system do the walking'. The system automatically records the user's activities and presents them symbolically in a virtual environment. The emphasis in symbolic acting is to show the contextual information telling us about where a user is, who they are and what they are doing right now, what documents they will use and what happened to documents.

Symbolic Acting takes away the responsibility from the user to navigate in a virtual environment or to control an avatar. The user is able to follow the events in the virtual environment as in a theatre, the system acts as a guide.

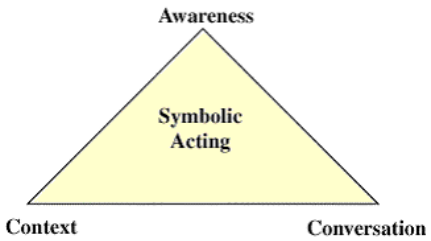


Figure 1: The Symbolic Acting Triangle

Symbolic acting enables the user to become aware of ongoing activities in the virtual team's work environment. It provides context on the activities of other users and encourages conversation on the current events (Figure 1).

The term Symbolic Acting was first used by BT in its project Forum [18] which aimed to show what each member of a connected group was doing by detecting their activities and representing these online in a clear, symbolic form. Symbolic Acting formed an important component in the project TOWER [24], where it was applied to present events in the TOWER virtual environment.

In the DocuDramas we employ Symbolic Acting in two different ways. At first we focus on symbolizing events in form of environmental cues. Thereafter we will introduce Symbolic Acting with avatars and present two DocuDramas based on this approach on Symbolic Acting. We will detail the employment of Symbolic Acting in the respective sections on DocuDrama.

3. THE DOCUDRAMAS

We present three different foci on DocuDrama, which represent user activity in a project workspace. The first DocuDrama symbolizes activities in their temporal structure in a time tunnel. The second focus lays on symbolizing the activities with respect to project context and goals to be met. The third focus takes the social perspective and symbolizes an activity as social relationship among the actors.

3.1 DocuDrama Timetunnel

In DocuDrama Timetunnel events are symbolized by means of environmental cues. To this end a space language provides a time-document-centred view in DocuDrama Timetunnel. The Timetunnel tells a story of the life cycle of a team's workspace. It visualizes folders and documents, deadlines and milestones. The aim is to provide an abstract view on related activities and to offer the functionality to manage data. The Timetunnel shows a symbolic space, the virtual representation of a project team's shared workspace (Figure 2). Moving through the tunnel enables a virtual journey through the project's lifetime, in which the tunnel symbolizes the time axis of the project.

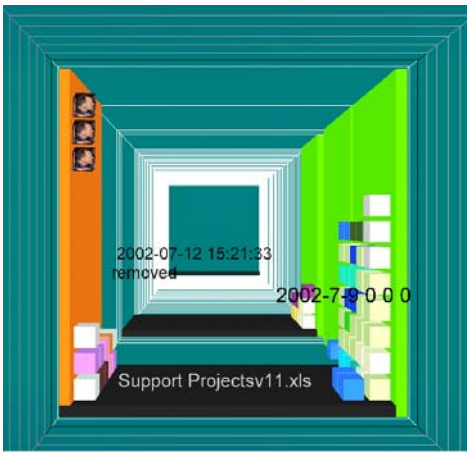


Figure 2: The Timetunnel

Small boxes placed on the wall of a time-slice symbolizes interaction with the project's folders and documents. Each box represents a document. The position of the box inside the time-slice indicates the form of interaction with the document. For example, boxes piled up on the right side of the time-slice might represent documents which have been opened for writing, boxes on the ceiling show documents which have been opened for reading. The colouring of the walls supports the meaning of the position.

Based on action events [24] over time a map of the folder's history is generated. Past events are aggregated in the form of a three dimensional environment, providing team-members with a

generative tool to visualize projects events history in various configurations.

In this prototype a 'meaningful' symbolic representation of events history is constructed by implementing a spatial approach that aggregates events and elements together in a chronological sequence as a configuration of related events.

In order to reveal the usually hidden relationships between separate strings of information, the events of a project are placed in a linear time-based spatial configuration, with the objective to create meaningful relationships between separate strings of events. The events that take place at the same time period will be interlinked and appear in the same time slice. In addition the documents could be coloured. This means that the box, which represents a specific document, will have a certain colour e.g red. It will keep the same colour, independent of the type of action performed on it. Its location inside the Timetunnel will be changed depending on the type of action performed on it.

Actions, events and milestones in the Timetunnel are arranged around the axis of movement forming the four surrounding walls. These walls have different colours depending on the kind of actions they represent.

Activities that have taken place on the same day are located in the same time slice. To display the information (date and the event type) that is represented by a specific wall, the user should click on the desired wall. The displayed events include read, delete, create, and move documents in a folder, which affect the state of that specific folder.

The evaluation in discussion with possible user groups showed that the DocuDrama Timetunnel represents an excellent tool for task control in case of shared document production. Users said that the three dimensional environment, which visualises space-time construction of event history, is easier to understand than pure textual list of events history. Instead the clustering of events is easier to detect. The user group pointed out that the Timetunnel could prove to be especially useful in relation to deadlines. This would enable the user to monitor the activities and delays in a task.

Future research and development will focus on the handling and visualisation of large datasets in the Timetunnel. To improve the functionality of the model we plan to experiment with context analysis and with different types of spatial clustering in a form of a parametric mapping of space-time configuration, which would reflect the actual number of events that took place at the specific time unit. Consequently the user would be able to detect and identify the period of high activity at a glance.

Future work as work as well will be to refine thee up to now quite complex symbolic representation. It will be enhanced with automatic focusing on areas of interesting activity, which both will simplify interaction with the Timetunnel and its content. The future Timetunnel will be useful to monitor the course of a task in relation to overall milestones and project deadlines. In particular it will offer functionality to manage and organize folders.

3.2 Symbolic Actions with avatars

In conversations between people, information is not only transferred by spoken words. Indeed language transmits only half of the message, the other and sometimes even more meaningful part of the message is conveyed through body language. Body language is the first language humans learn. Therefore humans are used to pick up information from the position, behaviour, and appearance of those around us [28]. Over the last decades Social Psychologists such as Michael Argyle [4] have done much to study the similarities and differences in which we react to crossed arms or a tapped nose or a shrug of the shoulders. Others, such as Desmond Morris [19; 20] have popularised these notions.

Results of these studies have been applied in the research on Embodied Conversational Agents [9]. This research direction focuses on the design of computer interfaces, with which the user can interact in a dialogue, under respect of conversational behaviour, emotion, personality and social convention. The interfaces have bodies, usually human-like, and perform movements of body language [3; 8; 15].

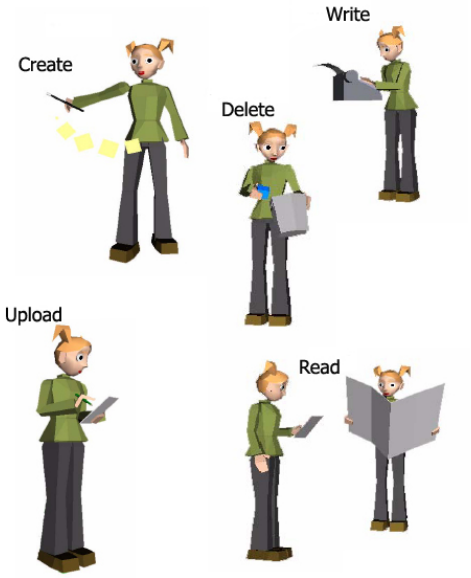


Figure 3: Symbolic actions of avatars

In the DocuDramas presented in the following individual team members are represented through avatars. The avatars perform symbolic actions which are derived from human vocabulary of body language. These gestures are intuitively understandable, enable to transmit the underlying information fast and effectively and at the same time convey a feeling of emotion and familiarity.

Figure 4: Avatar Customiser

In both approaches, activities of individual team members are represented through symbolic actions performed by avatars. For example, an avatar reading a newspapers symbolizes that the person behind the avatar is currently opening and reading documents. An avatar creating a series of stars with its wand symbolizes the creation of a document. Figure 3 shows a selection of the most commonly used Symbolic Actions.

The avatars are configurable and might therefore express the personality of their owners. The user can choose the avatar's hairstyle and colour, formal or casual clothes and two different options for a Symbolic Action. Figure 4 shows the Avatar

Customiser [21], which serves to personalize the users avatars. It has been developed in the course of the TOWER project.

3.3 DocuDrama Project

DocuDrama Project presents team-members with overview scenes of historic events in an axonometric environment. This is a pseudo 3D space in which team-members and projects are represented by 2D interactive icons (Figure 5). The arrangement of the space enables all avatars, and their contextual activity to be seen at a glance. The environment features both symbolic acting and visual annotation (paths) to display activity in the projects.

MobilesDisco, a chat space [22], uses a similar visual arrangement. It offers an axonometric view on a room in which avatars interact and communicate with each other. Some of the avatars represent people, others represent bots. The user can follow the activities in the space from an overview perspective and see his/her avatar interacting with other inhabitants of the MobilesDisco world. This relatively detached view enables the spectator to see the world more like a theatre and less than as a simulation of a real world [28].

Another inspirational source for the development of DocuDrama project originates from the game 'The Sims'[2]. In that game the player can look on and manipulate a whole neighbourhood of people. The axonometric view on the room perfectly enables the user to control all activities in the world and manipulate its inhabitants. A task which would be much more difficult to perform if the world would be represented by looking through the eyes of an individual.

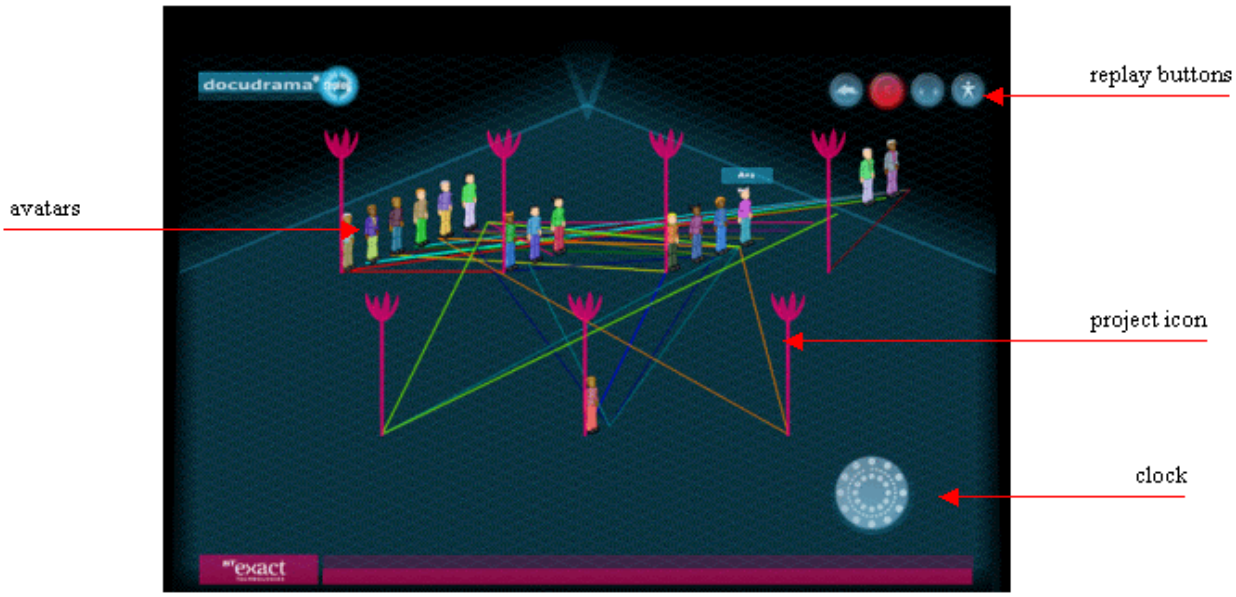


Figure 5 : Project Summary View

Built in Macromedia Flash MX, the interface runs on a client machine and can query local and shared project history files (in XML format). The interface has two parts, a configuration tool for project selection and a review interface. The latter presents the user with the options to refine the time period on display, and to view visual and textual summaries of events in that period. These summaries can be expanded into fully animated replays in which the avatars act out each event sequentially.

The symbolic actions used are renderings of a subset of the animations available in the original TOWER world, of which Read, Edit, Create and Delete are the most commonly detected. The animations involve the use of props to demonstrate the activity using commonly recognised icons – a waste bin for Delete, and a document for Edit. In the case of DocuDrama the avatar is seen to affect the prop in some way (throw the document in the waste basket). The avatars are rendered at a 45 degree angle, camera to the right and above the avatar, to define the most clear silhouette for each action.

The Project world is organised on a grid layout. According to the number of projects selected for review, the relevant number of project icons are automatically generated and arranged on the grid. The program lays out the icons evenly to give the maximum space for each project, and therefore there is no semantic meaning implied. The projects in this prototype are represented by 'flower' icons. The icons are sprites with animation properties. A

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

flower which is closed will represent a lack of recorded activity, a more open flower has a higher percentage of use.

In the summary view, avatars stand in their last recorded context , at the side of the relevant project icon. A uniquely coloured path is associated with each avatar and denotes a summary of where events took place for each individual in the selected time-period. As well as indicating where people have been most active – this annotation enables the user to see where there has been little or no activity. This view is intended as a prompt to drill down into further information cues: the replay of an individual history, or of the group as a whole.

When a replay is activated using the replay buttons, the avatars move from flower to flower and perform the action symbolising their activity within the project (Figure 6). Avatars line up in order of activity in a project. Additional properties i.e. the names of projects and people are available on mouse rollover. These can be clicked on to retrieve a detailed text summary of individual document actions, and statistics on overall project activity. A new time period can be selected using the clock in the bottom right of the screen. The clock incorporates month, day and hour in a circular interface, and has a progress bar to show how far through the replay the user is.

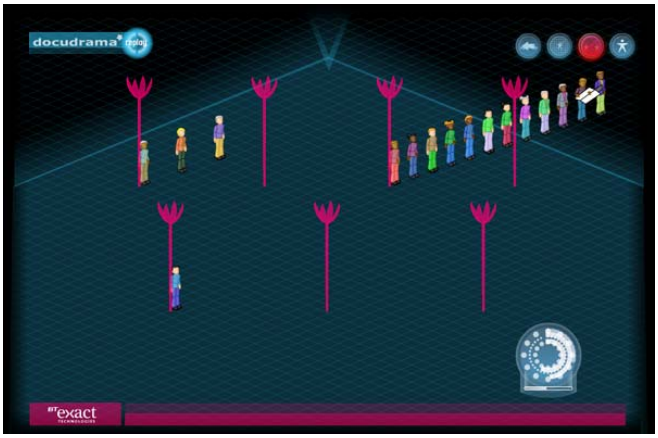


Figure 6: Replay View

The prototype realised has a limited set of features dictated by the data available for tracking events at the time of development. Integration into a project management tool which tracks deliverable status and team resources would be desirable. Feedback has been received from the TOWER user-testing group Atkins, and from teams experimenting with the interface in BT Exact, which includes project managers. The use of symbolic acting and annotated overviews were regarded as useful when compared to text based history files. We would seek to reduce the complexity of the animations and develop symbolic poses for speed of use and comprehension, and try to further understand the patterns of interaction over time, for example the intersection of paths could suggest a particular relationship between team-members on a piece of work. Users expressed a wish to interact directly with the progress bar to stop and start replays and to scrub backwards and forwards through time.

3.4 DocuDrama Conversation

In DocuDrama Conversation the social perspective is in the fore. Asynchronous interaction between people should be made visible. Therefore DocuDrama conversation focuses on the interaction between people on occurring on documents [27]. It uses spatial bodily positions and nonverbal communication to symbolize interaction between team members thus disclosing non-verbal communication sequences.

Film and theatre grammars [5; 28] also proved to be an effective resource of inspiration by the development of DocuDrama Conversation. The language of film and camera directing has evolved over the years with the audience. We are now at a point where some relatively subtle signals can convey precise ideas about everything from the weather and the passage of time to the innermost thoughts of a character. The film camera controls not only what we see but also the way how we understand it. A dialogue between two actors can appear like a normal conversation between two persons if regarded in a distance. With a camera close-up switching between the faces it can create a feeling of suspense and tension [12]. Camera controls not only what we see, but also how we see it. It determines our viewpoint, directs our intention, provides sympathy or antipathy to something /somebody, etc.

Certain similarity exists between film and computer graphic applications: both communicate a story mainly in images, both present a 3D world on 2D screen, both control the sensation of the audience by directing the camera[11]. In DocuDrama Conversation we employed film and narrative concepts of cinematography in the development of an automatic camera direction.

The story of the conversations on documents is presented in a three-dimensional virtual environment (Figure 7). The DocuDrama world symbolizes the shared workspace of a virtual team. The coloured and labelled boxes in the virtual environment denote different folders in that workspace. The boxes are coloured differently to symbolize the context of the folders. For example, the blue boxes in a row all belong to the context ‘Workpackages’. The tower in the background serves as landmark and, if the users wishes to explore the world by him/herself, as support for navigation [14].

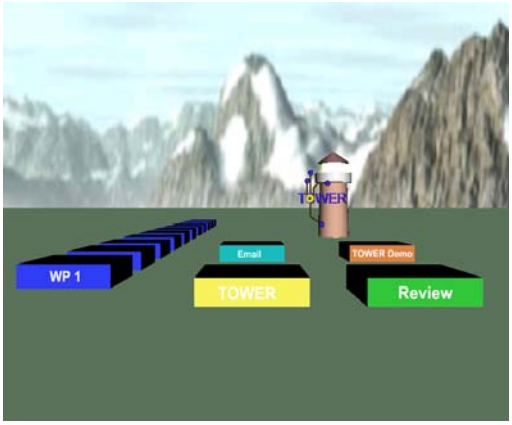


Figure 7: DocuDrama world

People acting on the same document in a given timeslot are positioned on top of the respective document. They are arranged

and filmed according to film idioms[5]. For example, two avatars interacting on the same document are positioned in one line and turned to each other (Figure 8). Several avatars interacting on the same document are arranged in a circle thus symbolizing potential collaboration (Figure 9).

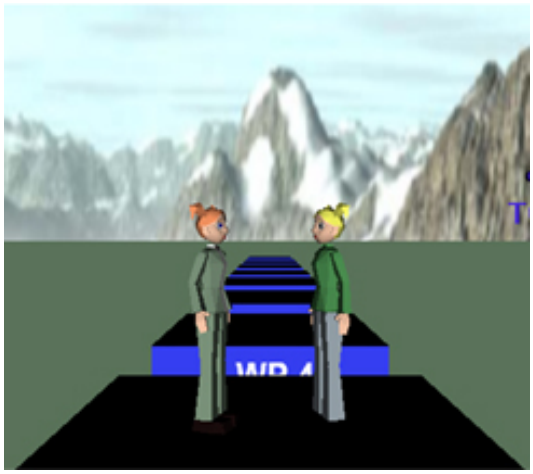


Figure 8: Avatars in a dialogue

Figure 9: Several avatars interacting on a document

The DocuDrama camera guides the user through the story of conversations. Cinematography rules are applied to present the story to the user in an entertaining way. The camera direction uses the position 'Medium shot', a camera position between overview on a scene and close-up on actors, in order to direct user attention to an action. It then moves forward to gives the spectator a closer look on the symbolic action performed by the avatar. The avatars appear one after the other and perform their action. If there are more then two avatars involved the camera shows their performances first before giving an overview on all participants of the conversation. The camera always moves at the same height as the avatars. Figure 10 shows how a dialog scene with three actors is filmed. First, an establishing overview shot is used to introduce the location to the user, at which the action is going to take place. Then the camera moves forward to show the individual symbolic actions by the different actors. At the end of the scene, a finishing overview shot is used to show at one glance which actors interacted with each other.

Improving the understanding of the collaboration processes is the aim of DocuDrama Conversation. The replay of events in DocuDrama conversation shows the team members' activities over a period of time. The story brings out the centre of interaction on documents in the teams' shared workspace as well as the sequence of interaction between the team members.

DocuDrama Conversation has been evaluated in two experiments with different settings. Both experiments focused on possibilities of a history replay of events. The user experiments have shown that it was easy to identify the most important documents in the work process. The users could easily point out the team members involved in the collaboration process and also the most active participants in the course of events. It was easy for the test users to identify the purpose of the collaboration, e.g. writing a paper because the sequence of events includes several Write, Create, and Read activities. In summary the test users liked the tool, although they criticised the slow motion of the story play-out and the sometimes tiring repetition of avatars performing Read-actions. Future Work will therefore focus on the development of a version of DocuDrama which offers a fast-forward overview on activities as well as a higher density of information.

4. CONCLUSION

In DocuDrama we introduced three different modi to represent project activities and interaction between team members. Each of the approaches uses a symbolic language of its own. They use different signs and symbols to represent project context and project-related events. The DocuDramas are designed to present a project in a specific point of view. Their representation and use of symbols responds to different sets of requirements defined by the user to fulfil a specific task, e.g. to visualize the project's workflow. Depending on the information needs a user may select the most appropriate symbolic visualisation. Although currently there are three different symbolic presentations, all of them are by nature of their visualisations easy to grasp and support intuitive understanding.

Future work will focus on the development of a one-for-all DocuDrama, which combines features of the DocuDrama approaches presented in this paper.

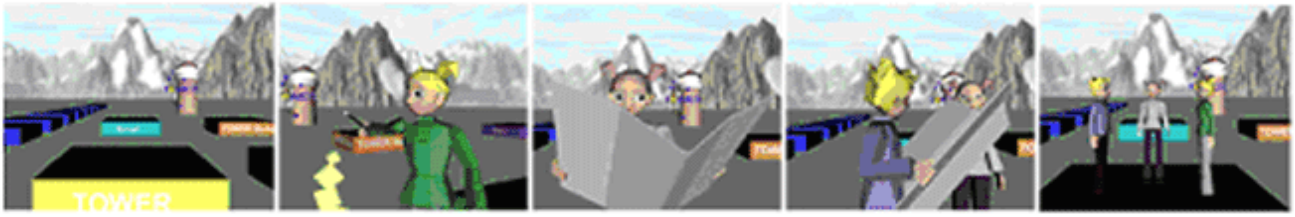


Figure 10: Dialog Scene with three actors

5. ACKNOWLEDGMENTS

We thank the members of the TOWER project team. The TOWER project has been partly funded through the IST program IST-10846.

6. REFERENCES

[1] Dramatica (2003), http://www.dramatica.com/T

[2] The Sims (2003), http://thesims.ea.com/

[3] André, E., et al. (2000). The Automated Design of Believable Dialogues for Animated Presentation Teams. Embodied Conversational Agents. J. S. J. Cassell, S. Prevost and E. Churchill, MIT Press: 220-255.

[4] Argyle, M. (1990). Bodily Communication. London, Rotledge.

[5] Arijon, D. (1991). Grammar of the Film Language. Los Angeles, Silman-James Press.

[6] Brooks, K. M. (1996). Do Story Agents Use Rocking Chairs. Fourth ACM international conference on Multimedia, Boston, USA, ACM Press.

[7] Brooks, K. M. (1999). Metalinear Cinematic Narrative: Theory, Process, and Tool, MIT.

[8] Cassell, J. (2000). "Embodied Conversational Interface Agents." Communications of the ACM 43(4): 70-78.

[9] Cassell, J., et al., Eds. (2000). Embodied Conversational Agents, MIT Press.

[10] Chatman, S. (1980). Story and Discourse: Narrative Structure in Fiction and Film, Cornell University Press.

[11] Clarke, A. and G. Mitchell (2001). Film and the Development of Interactive Narrative. ICVS 2001 (International Conference on Virtual Storytelling 2001).

[12] D.Katz, S. (1991). Film Directing Shot by Shot, Michael Wiese Productions.

[13] Fuchs, L., et al. (1995). Supporting Cooperative Awareness with Local Event Mechanisms: The GroupDesk System. Fourth European Conference on Computer-Supported Cooperative Work - ECSCW '95, Stockholm, Kluwer Academic Publishers, Dordrecht.

[14] G.Vinson, N. (1999). Design guidelines for landmarks to support navigation in virtual environments. CHI 99, Pittsburgh, PA USA, ACM Press.

[15] Gratch, J., et al. (2002). " Creating Interactive Virtual Humans: Some Assembly Required (Special issue on AI in Interactive Entertainment)." IEEE Intelligent Systems 17(4): 54-63.

[16] Greenhalgh, C., et al. (2000). Temporal Links: Recording and Replaying Virtual Environments. 8th ACM international conference on Multimedia (MM 2000), ACM Press.

[17] Gutwin, C., et al. (1996). Supporting Workspace Awareness in Groupware. Conference on Computer-Supported Cooperative Work - CSCW'96, Boston, MA, ACM, N.Y.

[18] McGrath, A. (1998). "The Forum." SIGGROUP Bulletin 9(3): 21 - 25.

[19] Morris, D. (1967). The Naked Ape, Vintage Paperbacks.

[20] Morris, D. (1997). Bodytalk. Körpersprache, Gesten und Gebärden. München, Wilhelm Heyne Verlag.

[21] Oldroyd, A. and J. Bulman (2001) Symbolic Acting Module D 5.2, BT, TOWER Report

[22] MobilesDisco (2000), www.mobilesdisco.com

[23] Prinz, W. (1999). NESSIE: An Awareness Environment for Cooperative Settings. Sixth European Conference on Computer Supported Cooperative Work (ECSCW '99), Kopenhagen, Kluwer Academic Publishers.

[24] Prinz, W., et al. (2003). Presenting activity information in an inhabited information space. Inhabited Information Spaces. E. Churchill and D. Snowdon.

[25] Prussak, L. (1997). Knowledge in Organizations. Oxford.

[26] Schäfer, L., et al. (2001). Virtual Storytelling of Collaborative Activities in a Theatre of Work. International Conference ICVS 2001: Virtual Storytelling, Avignon, France, Springer.

[27] Schäfer, L., et al. (2002). DocuDrama Conversations. SIGGRAPH 2002, San Antonio, Texas, USA, ACM.

[28] Smythe, P., et al. (2000) Symbolic Acting Module D 5.1, TOWER Project, TOWER Report

[29] Swan, J., et al. (1999). "Knowledge management and innovation: networks and networking." Journal of Knowledge Management 3(4): 262-275.

IDtension: The Simulation Of Narrative

Nicolas Szilas
Idtension
1, rue des Trois Couronnes
75011 Paris, France
+33 1 43 57 35 16
nicolas.szilas@libertysurf.fr
www.idtension.com

ABSTRACT
This paper describes a demonstration of a running system for Interactive Drama, called IDtension.

1. THEORETICAL FOUNDATIONS
IDtension is a 4 years old research project which aims at providing a new form of digital art, called Interactive Drama: this is drama on computers where the user participates in the story designed by the author by taking the role of one of the characters [6][7][8][9].

In order to solve the seemingly contradiction between interactivity and narrativity [5], our approach consists in simulating the narrative, and letting the user interact with this simulation. The simulation is based on two major components:

- A structural model of the narrative, that is an atemporal model of the narrative [8].
- A user model of narrative perception [7][9].

The user model, inspired by the idea of a "model of the reader" [2], allows to appropriately unfold on the fly the temporal sequence of the story from the structural model, according to the user's actions.

Thus, as discussed at COSIGN 2002 [8], structuralism plays a central role in the system: the structural description of the narrative, inspired from various structuralist theories [3][1][10], might be the only possibility to describe the story in such a way it can be really simulated (not juste reassembled)

2. OVERVIEW OF THE COMPUTER MODEL
A detailed description of the computer model can be found in [9]. We just provide here a flat list of the various type of data that are handled by the narrative engine:

- Character: entities which have goals and perform actions.
- Goals: Some states in the world of the story that characters want to reach.
- Tasks: Concrete activities which lead to the goal.
- Obstacles: Practical elements in the world of the story which make some tasks impossible.
- Actions: What characters do, including information transmission, influences, task performances and transformations.
- Character's states: Characters's wishes, knowledge, opportunities for action, etc.
- Values: Author-defined axes according to which the tasks are evaluated.
- Narrative effects: Criteria according to which a succession of actions is satisfying or not, from a narrative perspective.

The systems works at the level of the action: it takes as an input the action decided by the user, and it produces the next action in the drama.

3. WRITING A STORY WITH IDTENSION
Writing a story with IDtension consists mainly in building the structural model of the drama, which contains: characters, goals, tasks, obstacles and values, and the links between these elements. Such a structure can be represented as follows:

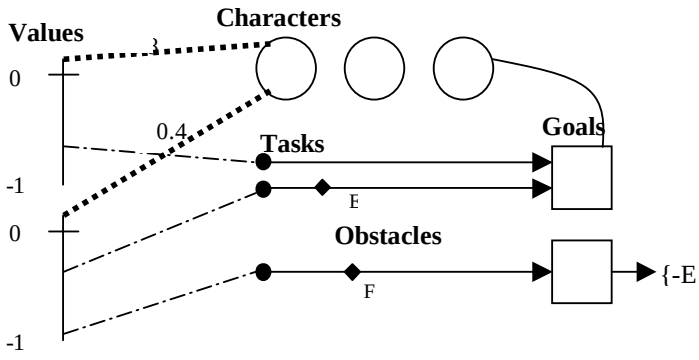


Figure 1. Structural description of a story. Characters (circles) wish to reach some goals (squares). Wishes are represented by curved lines. Each goal can be reached through tasks (arrows) that are more or less negatively evaluated according to each value of the narrative (dashed lines). The characters are more or less linked to the values (bold and dashed lines). Obstacles (diamonds) allow the triggering of sub-goals (via condition "E").

First published at COSIGN-2003,
09 – 12 September 2003, University of Teesside (UK),
School of Computing and Mathematics, Virtual Environments Group

Such a structure describes the narrative in an atemporal manner. The narrative engines uses this structure to unfold a meaningful story, in an interactive manner.

4. THE TECHNICAL DEMONSTRATION

The IDtension system is fully developed in Java. All modules are operational, even if each of them is under continuous development.

In the current version, at the time of writing, the interface is reduced to its minimal form: actions are chosen by the user among a list of all possible actions, and actions are displayed in a textual and predicative form.

Depending on the narrative mode chosen for the narrative sequencer, there are several ways to activate the system. Currently, we have two modes:

- automatic generation: the system chooses one action among the best actions, which is sent to the user;
- first person: the user is responsible for all of the actions of one character. The user and the computer alternate their action, like in a chess play.

A better interface is currently under development.

Two scenarios have been developed. The first one is very simple: it involves 6 characters, 3 goals, 4 tasks, 3 obstacles and 2 values (see [9]). A new scenario, involving more goals and more tasks is under development, and will be released for the demonstration.

For this second scenario, the rendering of actions is performed through an simple natural generation system: from pieces of sentences written by an author, it generates a natural language form for each action in the story. We will demonstrate how this technical choice is a good compromise between generativity and ability of authoring. Currently, the language of the output is French.

5. REFERENCES

[1] Bremond, C. Logique du récit. Seuil, 1974.

[2] Eco, U. Lector in Fabula. Bompiani, 1979.

[3] Levi-Strauss, C. Anthropologie Structurale. Plon, 1958.

[4] Murray J. Hamlet on the Holodeck. The future of narrative in the cyberspace. Free Press, 1997.

[5] Ryan, M.-L. Narrative as Virtual Reality. The John Hopkins University Press, Baltimore, Maryland, USA (2001)

[6] Szilas, N. Interactive Drama on Computer: Beyond Linear Narrative. In Proc. AAAI Fall Symposium on Narrative Intelligence (North Falmouth MA, November 1999), AAAI Press, 150-156. (also available at http://nicolas.szilas.free.fr/research/aaai99.html)

[7] Szilas, N. A New Approach to Interactive Drama: From Intelligent Characters to an Intelligent Virtual Narrator. In Proc. of the Spring Symposium on Artificial Intelligence and Interactive Entertainment (Stanford CA, March 2001), AAAI Press, 72-76. (also available at http://nicolas.szilas.free.fr/research/aaai01.html)

[8] Szilas, N.: Structural Models for Interactive Drama. In the Proc of the 2nd International Conference on Computational Semiotics for Games and New Media (Augsburg, Germany, Sept. 2002)

[9] Szilas, N. IDtension: a narrative engine for Interactive Drama. In Proc. of TIDSE'03 (Darmstadt, Germany , March 2003), AAAI Press, 72-76. (also available at www.idtension.com)

[10] Todorov, T. Les transformations narratives. Poétiques, 3 (1970), 322-333.