71

Animist Interface:
Experiments in Mapping Character Animation to Computer Interface

by

Richard W. Lachman

B.S. Computer Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA
1995

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE
IN MEDIA ARTS AND SCIENCES
at the
Massachusetts Institute of Technology
September 1997

Signature of Author ______________________ 8/8/97
Program in Media Arts and Sciences
August 8, 1997

Certified by ______________________ 8/8/97
Glorianna Davenport
Principal Research Associate
Media Laboratory
Thesis Supervisor

Accepted by ______________________
Stephen A. Benton
Chairman
Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Animist Interface:
Experiments in Mapping Character Animation to Computer Interface

by
Richard W. Lachman

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on August 9,1997,
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN
MEDIA ARTS AND SCIENCES
at the
Massachusetts Institute of Technology
September 1997

Abstract
This thesis presents a collection of techniques for integrating character and setting with software agent communication. "Animist Interface" supplements and supplants traditional interface elements such as plain-text dialogue boxes, static images and monotone audio to create a richer, more cinematic messaging environment. The techniques animate subtle nuances on the periphery of the user's attention to increase the bandwidth of computer/user communication.

The problem area is defined by two questions: First, how can we use character-based interfaces to make agent communication more efficient, meaningful, and enjoyable? Second, how can cinematographic techniques for directing, supporting and presenting character and setting inform this interface? Several principles of Animist Interface are explored, including the use of lightweight characters for communication, the scaling of prominence with relevance, and the use of virtual lighting and camera-work within an interface.

Simple software agents were created, which communicate proactively with the user in the form of characters or through cinematic techniques. Three "functional sketches" were used to put these ideas into practice: 1. An anthropomorphic interface; 2. A system based on background video imagery; 3. a project that combines virtual camera techniques, lighting and simple 3-D characters. Further directions and an outline for a future project/user-study conclude the thesis.

Thesis Supervisor: Glorianna Davenport
Title: Principal Research Associate

Animist Interface:
Experiments in Mapping Character Animation to Computer Interface

by
Richard W. Lachman

The following people served as readers for this thesis:

Reader: ______________________________

Pattie Maes
Associate Professor
Sony Corporation Career Development Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Reader: ______________________________

Bruce M. Blumberg
Assistant Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Table of Contents

chapter one

Introduction

Dealing with Overload

In the course of a day, we take in and process an incredible number of messages. A simple walk down the street bombards us with information from all sides, and each burst of colour, of sound, of smell speaks volumes about where we are or what is happening around us. Even with media-representations, when several channels of input are removed, we can still take in enough to follow the complex images, sounds and plot lines of a movie. In the middle of all of this noise, we don't complain about "information overload", or feel adrift in a sea of unfiltered knowledge.

The Internet and our off-line computer forays, on the other hand, are described as "drowning…in the rising tide of information"[Denning82], or wading through "heaps of dirt that [are] shipped to you on the infobahn everyday" [Lanier95]. Intelligent Agents, software designed to alleviate this overload, are described as "a deception and an empty promise" [Schneiderman96]. Why the massive disparity in our feelings of comfort and control?

The basic fact is that we communicate with the world around us in a variety of ways. Understanding a real world "message" involves many senses, at many priorities or levels of conscious thought. We do the "filtering" ourselves, since the way the message is presented says something about the message itself. Of course, there are incredible differences between these two situations: our physical presence among the "emitters of information" (the bus screaming around the corner, the honeysuckle along the hedgerow) gives us spatialization, a kinaesthetic understanding, a tactile relationship...a level of immersion which we, in developing technology, are just taking our first steps towards understanding.

The line between computer-mediated experiences and film is less clear. In many ways, the palette of technologies available to each seem the same: an interplay of light and sound, projected on a two-dimensional surface over time. However, our techniques for dealing with information overload within the computer world are woefully constrained. As a colleague remarked, "graphic design was a great resource for the 2-D, text-based computers of the past. Nowadays, we have fast 3-D graphics cards, stereo sound-systems, real-time animation..." [Wexelblatt97]. In other words, we have many of the same fundamental building blocks as the cinematographer, and we can use these to help communicate to the user in a new fashion. As Joe Bates notes, "The arts (film, theatre, novels, radio drama) have long studied a related problem [to user-interface and interactive entertainment agents]. Here the agents are called 'characters'...we believe these artists have much to convey to AI researchers trying to construct believable interactive creatures" [Bates94b].

This thesis explores new techniques of embodying information: using computer interfaces which draw upon what we already know about the language of cinema and the reactions of characters.

Fiction

In his novel "Idoru", William Gibson presents the following scenario within a virtual environment:

> 'The lizard she'd adjusted earlier came scurrying low across a narrow ledge of rock, clinging so close as to appear two-dimensional. Zona picked it up and stroked it into yet another color-configuration.
>
> "What are you doing?"
>
> "Harder encryption," Zona said, and put the lizard on the lapel of her jacket, where it clung like a brooch, its eyes tiny spheres of onyx. "Someone is looking for you. Probably they've already found you. We must try to insure that our conversation is secure."
>
> "Can you do that, with him?" The lizard's head moved.

"Maybe. He's new. But those are better." She pointed up with the stick. Chia squinted into the evening sky, dark cloud tinted with streaks of sunset pink. She though she saw a sweep of wings, so high. Two things flying. Big. Not planes. But then they were gone.

****

"I have some ghost-bastard thing on my site…"

"It's in the software on my Sandbenders, too," Chia said.

…

"Show me," Zona said. "In my site I could not see it. My lizards could not see it either, but they grew agitated. The birds flew lower, but could find nothing. Show me this thing!" [Gibson96].

The lifeforms within Gibson's virtual world communicate information through their behaviours and emotional states. Rather than explicit dialog-box interaction, the scene itself implies the information. At all times, "the computer" in the monolithic sense communicates through many embodied agents – however, the subdivisions and separate programs allow users to focus on different aspects, to compartmentalize tasks, to work in concert with the semi-autonomous software. And, through changes in lighting and atmosphere – common cinematic techniques – the emotional meaning of a message becomes clear. The goal is "to condense fact from the vapor of nuance," as a character describes a virtual environment in Neal Stephenson's Snow Crash [Stephenson93].

Why Animism?

The name for this system grows from the idea that information can be embodied and communicated through the very environment in which we interact. Animism is a principle expressed in many first-peoples religions around the world, with echoes in modern Japanese Shinto beliefs:

"[Animism] is the system of thinking based on the idea that the spirit (anima) is the seat of life and to [sic] explain all the phenomena of nature by attributing them to spiritual agency. It is common among primitive races, as field works of cultural anthropologists' report, and among infants in the early stage of learning logical thinking. It was probably the most primordial system of thinking for understanding the world by the human mind.

In animism, everything is supposed to have soul and spirit and act accordingly, i.e. "personification" of the components of the world. In the animistic world view, human being is not the supreme ruler of the world, but a member of the community with other beings, including animals, plants, minerals and natural phenomena such as the sky, water, mountains and the earth" [ICT95].

In the computer world, there is definite resistance to treating software as overtly animate. Dijkstra warns that anthropomorphism (specifically) results in a waste of mental effort [Dijkstra90]. Shneiderman reports that it puts the user at the mercy of the machine, instead of letting them feel like they can initiate and control actions [Shneiderman95]. The thought is seen as a mistake, a mental slip made when we aren't paying enough attention to our reasoning. This thesis, however, will attempt to show a beneficial side to this "slip". Animism, an "illusion of life", can serve to be a useful way for a user to work with the increasingly pro-active programs of the modern world. Nass and Reeves find that we already respond to computers according to fundamentally social rules [Reeves96]. By following ideas that seem to be more in line with the way we ordinarily make sense of the world, we can create a more fun, engaging and sensorally rich interface to our computers.

"Rationalists will say that this amounts to a mystification. Certainly, when you would look inside the casings of...machines or print out the code you would find a completely determined technical system. But for the perception there remains something that is not identical with the technical description of its functional parts. From an engineering point of view these are new interfaces. From an epistemological point they are attempts to animate our environment, make it respond 'as if' it understood our wishes."
[Grassmuck93] (sic) (note: translated from the German by way of Japanese)

Overview of Thesis

This thesis is an exploration of how to map cinematic techniques and character animation to interface. It consists of several "working sketches", which have lead to a more organized set of ideas. In the interests of clarity, these "theories" or directions of focus are introduced before the projects that motivated them.

Chapter 1, the introduction, serves to frame the general area, and to present a fictionalized view that inspired my thoughts.

Chapter 2 outlines the groundwork this thesis builds upon. Ideas from software agents, agents for entertainment, and theories of communication/theatre are discussed, using specific projects and papers from the literature.

Chapter 3 contains a more thorough discussion of the major ideas proposed by this thesis. I discuss the differences between animism and anthropomorphism, the nature of background communication, how to degrade presence when an agent is no longer relevant, lighting/camera techniques, and the importance of sound.

Chapter 4 describes the actual 'functional sketches'. I created a series of agents under the Microsoft Agent architecture, as well as two projects running video and 3-D graphics as part of a Windows NT desktop environment. Details of implementation and an analysis of strengths and weaknesses are given for each experiment.

Chapter 5 concludes the thesis with a discussion of future directions for research, and an outline for a proposed user-study.

chapter two

Prior Work

This thesis attempts to bring together ideas from three broad areas of research. To truly meet the goal of a more enjoyable and expressive interface, ideas from computer-science, social science and the arts can all be brought into the same arena. We are concerned with an intelligence that originates in the computer-science world, where there have been many relevant projects that deal with representing and conveying information as well as producing it. We can also gain insight from computer scientists who focus on how to communicate emotion through sound and image – specifically, agents created with an eye towards entertainment. Finally, a more analytic approach to the question of human comfort and communication can be gained by studying theories of human-human and human-computer interactions.

In this chapter, I discuss relevant points from a variety of existing literature, projects and products. The first main area, Personified Software Agents, concerns projects which have the intelligence behind software as their primary focus; these works bring to light several ideas relevant to our explorations of agents-as-interface. The second section deals with characters and agents produced for Entertainment, focusing on believable characters and human reactions to them; Finally, we look at a few theories of Communication, paying specific attention to ways in which social dynamics and the dramatic arts influence the way we deal with computers.

Personified Software Agents

The dream of an intelligent software agent is pervasive and long-standing. In the time from Selfridge's "demons" of the 1950's [Selfridge59] to Negroponte's famous wink to his wife [Negroponte96] to the commercial availability of agents like Firefly [Firefly], there has been a great deal of experimentation in learning strategies, action-selection schemas and user-modeling techniques. The goal is to create software that can be pro-active, assisting intimately and resourcefully as we interact with the computer. In this section I examine a few projects which express characteristic approaches to agent-representation, regardless of specific functionality.

Again, this thesis does not directly try to increase the intelligence of software, but to help its ability to convey information to the user. To that end, I will examine Apple Computer's "Knowledge Navigator" vision-video and "Guides" project, several projects from the MIT Media Lab in which facial representations are used to support agent-intelligence, and Microsoft's "Office Assistant", "Bob" and "Persona" initiatives.

Knowledge Navigator [Apple87]

Despite "experimentation" in the science fiction world for decades, Apple's Knowledge Navigator popularized the idea of the agent-as-eager-assistant. Apple produced a video envisioning a possible agent of the future, and later created a minimal software implementation to explore related ideas. The scenario showed a fictionalized interaction between a university professor and Phil, the archetypal "digital butler". The character, rendered as a photo-realistic human, provided laconic commentary as it dealt with the user's professional and private life with intimate familiarity.

The Knowledge Navigator seems to crystallize many of the problems and dreams surrounding anthropomorphized interfaces. To critics, the cold, smug, somehow Vulcan Phil represents a "superior" feel to the interface, putting the human in a more subservient role. The character dispassionately displays the progress of devastation in the rainforest [Maes97], and lies to the user's mother [Sterling95], conjuring images of an amoral intelligence within the machine. What's more, the character in the video doesn't capitalize on its visual representation at all: gaze direction, the expressions of the character, and even its tone of voice rarely vary. The natural-language system and underlying intelligence depicted seem to be much more important to the design, while the representation seems superfluous. Even the specifics of Phil's representation "(high-resolution, real-time human portrayal) [don't] match his language capabilities, his thoughts, or his actions (single tasks performed in a rather unimaginative manner)" [Laurel 1991]. Still others criticize the domain in which the agent was situated: another stultified office-automation product, working to the strains of the Masterpiece Theatre soundtrack. Indeed, later videos such as Apple's "Chapter One" envision the use of agent technology as useful in more naturalistic settings such as a home school, a writer's den or the kitchen. [Apple88]

To proponents of embodied agents, however, Phil represents a dream of what might be possible – as useful as a long-time human assistant, with persistent, intimate and multi-domain knowledge of the user. Phil would be sensitive to contexts, and could proactively suggest appropriate content. And while not exactly friendly, at lease he could respond to commands as a conversational participant.

To my mind, the project was extremely useful but represents a misguided view. The representation of Phil falls for the idea that a computer-generated person *is* what an agent should be. His embodiment isn't a good match for his abilities. Computer generated "people" will continue to receive an incredible focus of attention, as the idea has stimulated computer-research for decades in the past. However, with respect to computer-interface, it seems worthwhile to consider other embodiments that might be more useful.

Guides

The Apple Guides project used video characters as part of an encyclopedia of American History. Different characters (a frontiersman, a native american and a settler-woman) presented different perspectives on events. The system selected from various clips, recommending interpretations, topics and narratives to the user, in an effort to use the "biased" nature of such information in a positive and meaningful manner. One of the goals of the project was to connect "the story being told (content) and the conditions of its telling (structure and context) [Don90]. Although the characters seem more closely related to oral, theatrical, and filmic representations than to dynamically generated agents, the project is still quite interesting to our discussion. The anthropomorphic, story-centered characters were well received by users browsing the content, but misinterpreted by those searching for a specific piece of information. Laurel notes frustration on the part of users who wanted to give natural language conversational instructions to the guides, a level of sophistication being far beyond the capabilities of the agents [Laurel91]. It seems probable that the photorealistic, human character representation increased these expectations. The production qualities and rendering style created a mismatch, then, between the intelligence of the agent and the expectations of the user.

Agents with Faces

Several projects have created agents that communicate through face-only representations. An email agent ("Maxims") and a calendar agent produced at the MIT used a caricature to indicate nine distinct states for the agent. The face conveyed that the agent was "Working", or "Confused" or "Unsure" as it recommended actions to the user. The expressions were discrete, static and relatively simple, but made steps towards a feedback many agent-systems completely ignore [Lashkari94][Kozierok93].

Koda studied rendering-styles, producing a range of cartoon and realistic faces in the context of a poker game. The users played against various characters, qualities like apparent intelligence, expressiveness, ability and their own sense of comfort. Her study concluded that faces produce a greater sense of engagement and emotional attachment in users, but also that they are a distraction from the task at hand – in this case, the actual poker game. [Koda96]

Microsoft Bob [Microsoft1]

"Bob" was designed to be an animated, easy-to-use front-end for a set of home-computer applications (datebook, financial planner, etc). Different family members were invited to pick a personal assistant from one of several characters, each with its own "personality" (read: verbosity). The characters (cartoon-style dogs, cats, bunnies and the like) communicate through comic-book speech balloons, presenting the user with hot-linked options to help them learn about the software and execute macros.

Bob was the focus much criticism from users, computing magazines and the academic community, as it seemed to present a "dumbed-down" environment in which the computer was completely in charge. The product has been described as "propos[ing] to the user a life of caricatured meaninglessness, sliding unintentionally into the grotesque", "something straight out of Diane Arbus, " [Lanier95]. Bob is seen as a sickly-sweet, crippling use of character in interface. Despite the fact that "User studies with Bob have verified that for many people, the social metaphor reduces the anxiety associated with computer use" [Ball97], articles and user-reviews almost universally panned the system [Lettice95][Wohl95].

The key failure seems to be that Bob completely replaced familiar tools and software representations with "cutesy", overly representational ones. The user was thrust into an environment where all control seemed to need clearance by the software agents, with little to no direct manipulation possible at all.

Office Assistant [Microsoft97]

One of the more interesting productized systems was produced for the Microsoft Office '97 suite. The user can select an animated character – a Paper Clip, a cartoon Einstein, a bouncy red ball, etc – to be the source for all help-file accesses. In addition, the character can animate to show the machine following user-instructions (ie: printing, saving, sending email), and pro-actively suggests tips and shortcuts. Apart from well-crafted intelligence, the character scales its behaviour to the type of information it has to dispense: if its giving an ordinary shortcut tip, a light bulb appears next to it. If it is responding to a more important action (ie: the closing of an unsaved file) it taps at the screen and gesticulates wildly. For the most part, however, its animations don't vary – it doesn't give any more information about computer-responses to user actions than a simple acknowledging of commands. In fact, most of the time the characters aren't conveying anything and the user has the option iconizing the figure until it's wanted again.

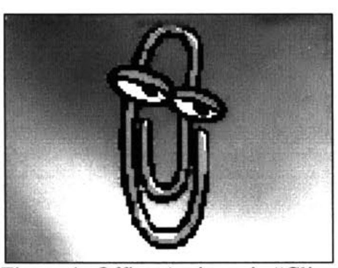

Figure 1: Office Assistant's "Clippit"

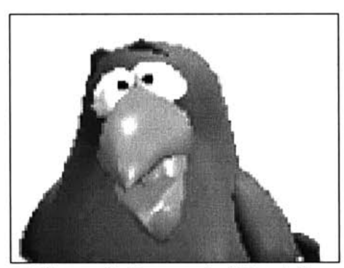

Figure 2: Persona's "Peedy"

Persona [Ball97]

The Microsoft "Persona" project is a more recent example of work in the field. "Peedy", a computer-generated parrot serves as the interface to a music-selection/recommendation system. The project focuses on conversational interface, the goal being a character that can react to and respond with spoken English in a coherent dialogue. Peedy is rendered as a 3-D computer-graphics character, with synchronized movements and speech output. When asked, "What have you got by Bonnie Raitt", Peedy waves in a stream of notes and grabs one as they rush by, saying "I have 'The Bonnie Raitt Collection' from 1990". Peedy keeps track of which conversational state, allowing the user to respond with "Pick something from that". The animation engine supports the dialogue, for example having Peedy cup his wing to his ear while asking the user to repeat themselves when a command isn't understood.

Future goals of the system include plans for intelligent virtual-camera control (ie: to automatically track moving objects), and a tighter relation between the animations and the character (ie: the addition of emotional state and a more meaningful gestural language to Peedy's animation-repertoire).

Peedy presents a well-developed, consistent character *as* the interface, rather than an animated icon serving to indicate state for intelligence which is clearly positioned elsewhere. Still, the system does focus on the assistant as ever present and always the center of attention. In effect, Peedy is a direct-manipulation agent -- if the mixing of terminology doesn't make that description too confusing. Peedy is more like the librarian in a private, restricted-access collection than a helpful sidekick suggesting information to you as you go through daily life. This is precisely the philosophy Shneiderman predicts will prevent agent-style interactions from gaining commercial success: the agent always is in control, and user has no sense of mastery over the actions [Shneiderman94].

Entertainment

The world of entertainment raises many interesting questions not addressed in debates surrounding task-oriented software-agents. Indeed, "…building entertaining agents requires the agent researcher to think more about the user. The researcher is forced to address the psychology of the user; how will the typical user perceive the virtual characters, what behavior will s/he engage in, what misconceptions and confusing situation may arise, and so on" [Maes95].

Many of the following projects have been influenced by traditional animators, most notably the Disney philosophies, as espoused in "Disney Animation: The Illusion of Life" by Frank Thomas and Ollie Johnston. Guiding principles of animation ("Do not let the expression conflict with the dialogue. The way a character walks, stands, sits, listens – all reveal the meaning of his words"), characterization ("From the earliest days, it has been the portrayal of emotions that has given the Disney characters the illusion of life.") and even choice of rendering styles ("Tell your story through the broad cartoon characters rather than the "straight" ones. There is no way to animate strong-enough attitudes, feelings, or expressions on realistic characters to get the communication you should have [sic]. The more real, the less latitude for clear communication") [Thomas91].

In this section I will examine the Oz project at CMU, the Improv project from NYU, Blumberg's Hamsterdam Construction Kit and Silas character, the P.F. Magic Petz products, Fujitsu's new "FinFin" environment and "Creatures" from CyberLife.

Oz [Bates94]

Joe Bates and his students at CMU have produced a project called Oz, developing interactive characters with a coherent emotional subsystem. The Woggles, who populate the world, are 3-D rendered autonomous characters existing in a cooperative environment, interacting with one another, their world, and the user (through simple mouse-clicks). The project has served as a testbed for a number of techniques in what Bates refers to as "believable characters". The goal is to produce creatures with such a consistent, predictable, dynamic set of behaviours that they appear to be reacting to stimuli around them in a life-like fashion. The main emphasis of the project is on making well-expressed emotional states, as a basis for more complex behaviours and goal-oriented actions. However, by focussing only on the characters, they seem to have given themselves a much more difficult problem. One of his students notes he "started with what I thought would be a reasonable looking behavior, but it wasn't clear enough so I tripled and then doubled again [the relevant parameters] and still only got a mope that people recognize when its pointed out to them." In addition to emphasized behaviours, animators and filmmakers alike use the scenery and lighting of a shot to provide emotional reinforcement. Oz seems a likely candidate to benefit from the ideas developed in the next few chapters.

Improv [Perlin96]

Ken Perlin and Athomas Goldberg at NYU have been working on characters capable of semi-autonomous action, at many levels of control. The basic design-impetus has been to create characters that can serve as "digital extras" in movies, interactive stories and environments of the future. The "actors" must be able to respond to commands from a plot, from internal goals/behaviours, or from direct user-manipulation, integrating all input smoothly and seamlessly. Improv characters incorporate noise functions as part of their life-like animation system, ensuring that their actions are never repetitive or mechanistic. Characters walk with a subtle, varying sway, or hold their heads at a slight tilt, or animate any of a variety of other degrees of freedom designed to make them more naturalistic. Furthermore, they react to other characters, and are able to create coherent scenes together with minimal direct authorial control. A key lesson learned form their work is that small motions in the animation can make a big difference in how the character is perceived. An audience picks up on tiny, almost imperceptible details in character-animation, implying we can use and vary those tiny details for different effects.

Hamsterdam/Silas [Blumberg96]

At the MIT Media Lab, Bruce Blumberg's "Hamsterdam Construction Kit", and Silas T. Dog (the flagship character created with it) uses an ethological basis for behaviours and renders them with attention to tips from the Disney animators. Competing goals within the character respond to stimuli, with a particular behaviour eventually "winning" (and being expressed by the animation/motor-control subsystems). However, the characters are still affected by the secondary, "losing" behaviours, which may control other facets of the character's appearance. For example, if Silas wants to play but the user orders him away, the gross motor action is to make the dog leave. Secondary controls, however, may make his head look back at the user, forlornly drooping his ears and tail. In this way, emotions seem to come alive in the character. Rather than a simple collection of finite states, the creature gains expressive power through its ability to display many, perhaps even conflicting, states at the same time.

Petz

PFMagic's "Interactive Pets" technology has produced the most widely distributed and well-received examples of interactive character technology to date. Using simple behavioural and learning algorithms, the Dogz, Catz and Oddballz characters run across their owners computer-desktops, playing, sleeping, and performing tricks at their owner's encouragement. People develop strong feelings of ownership and affection for the digital pets, as evinced by the database of "Petz snapshots" submitted by proud owners, or the raging debate on "Should Virtual Petz Die" on the PFMagic website [PFMagic]. The characters are reasonably simplistic in animation-style and movement (existing as 2.5-D animations in common Windows or Macintosh desktop environments), yet produce easily recognizable and engaging behaviours.

FinFin

Fujitsu Interactive has worked with Joe Bates to release the Teo world – a PC-based virtual environment, populated by various "alien lifeforms" and available for consumer purchase. The first of these characters has been released to market, and presents as a 3-D mutant dolphin named FinFin. The character exists within a graphically rich, rendered screen-world, going about his daily routines of eating, playing, singing, etc. The user can interact with FinFin by attracting him with virtual food, or through the "Teo Antenna", an external motion-sensor and microphone plugged in to a computer through the game-port. Over time, FinFin will become "accustomed" to the human visitor, and may even perform tricks, or learn to come closer when a whistle is blown. The character is placed within a coherent environment which parallels our own: Planet Teo is synched to earth time via the system-clock, and FinFin's behaviors, moods, and activities are influenced by the rising and setting sun. The "supporting cast" of plants and animals adds to the believability of the main character, and helps to convey his emotional state. FinFin singing as the sun sets behind him is made much more beautiful by the dynamic response of the environment to the character (or vice versa). Unfortunately, the creators have fallen back on traditional user-interface devices in this Alife project: rather than relying on FinFin's acting abilities, "users can monitor FinFin's feelings via a 'biosensor,' which indicates the level of happiness, fear, hunger, thirst, fatigue, sleepiness, etc., he is experiencing" [Fujitsu97]. The biosensor is a pop-up status-bar, with graphs and indicators to reveal something that really should be apparent through the actions of the character itself.

Creatures

"Creatures" is another Alife game/environment, where cartoon "Norns" interact, grow and learn in an on-screen world. Users can communicate with the characters through simple text commands or with a mouse-pointer, and can observe them with a first-person point-of-view (to see what a particular creatures is looking at), a third-person view as they learn about and

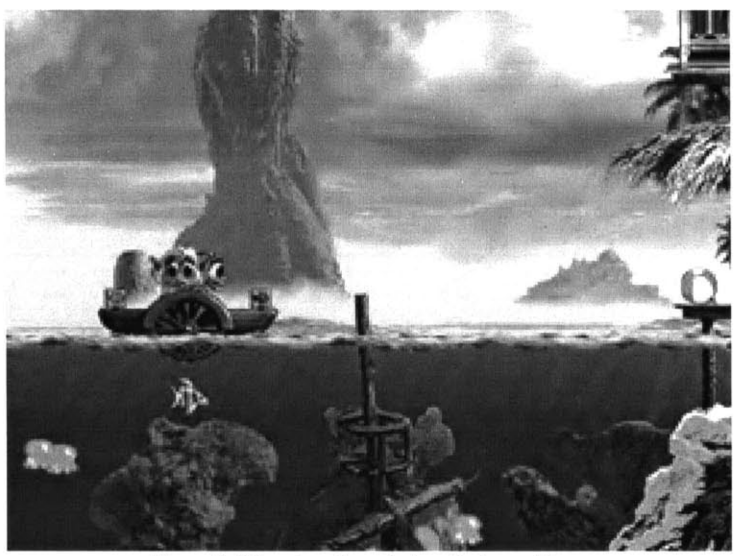

Figure 3: "Creatures" exploring their world

interact with their world, and a pop-up window to inspect personal and genetic information for a creature. Users can trade Norns on the Net, breed them to emphasize certain traits, teach them words and concepts, and let them teach one another about the world they live in. The ecosystem in Creatures is quite complex, with various plants, diseases, and predators one must help the Norns deal with. As the User's Guide mentions, "Please note that the Norns look to you as their keeper and parent," [CyberLife97] and the creatures can develop bad habits, diseases or poor genetic diversity through "bad parenting". A cross between SimCity™ and Petz, the software is a great vision of a complex, interactive, character-based ecosystem. It would be interesting to have on-screen variables like weather map state to real-world or on-line activity, with the health and behaviour of the Norns communicating information as well as entertaining.

Communications/Theories

This work also depends on certain ideas about the way peoples' interaction with machines and with one another. 'Animist Interface' hinges on electronic media's ability to influence us subtly and subconsciously with common and relatively simple techniques. Everything from the style of presentation (ie: film techniques and dramatic sensibilities) to basic social interactions can be used to influence users, implying information without needing full text. Nass and Reeve's studies of the parallels between human-computer and human-human interactions, Brenda Laurel's theories from her book "Computers As Theatre" and Clanton's contrasting of Film and UI techniques are discussed below.

The Media Equation [Reeves96]

Professors' Nass and Reeves of Stanford's "Social Responses to Communication Technologies" center have re-created more 50 experiments from social-science literature, designed to test the specifics of our social interactions. However, in each case, they modified the protocols to test if the same responses could be elicited from human-computer relations. Overwhelmingly, the results detailed in their book "The Media Equation" lead to the conclusion that our computer/media interactions are fundamentally social. Dominant personality types, team-player scenarios, praise/criticism situations, even notions of politeness all come into play when people are dealing with computer interfaces. The findings don't suggest that we think of computers as people: rather, that we perceive many social qualities in our interactions with machines, and are affected by them in the same ways as we are by people.

While their philosophies and ideas have been very influential to my thinking on interfaces in general, the following ideas are especially relevant to our discussion:

-The relative importance of rendering styles and "realism" in anthropomorphic interface is not as great a concern as commonly thought. "Simple line drawings of objects (as long as they have eyes or even a hint that they are alive", are quite enough to activate psychologically rich responses."

-People can identify personality (predict reactions, characterize affinities) in almost any type of interface, "from toaster ovens and televisions to word processors and workstations. Personality can creep in everywhere – the language in error messages, user prompts, methods for navigating options, and even choices of type font and layout."

-'Teammate' is a better way to approach human-computer interaction than either 'tool' or 'all-knowing agent'. "Computers and users should be peers, and people should be made to feel dependent on computers without feeling inferior or superior. Being on the same team encourages people to think that the computer is more likable and effective, and it also promotes cooperation and better performance"

The philosophy of the Media Equation isn't limited to character-based, humanoid interfaces, but discusses the effects of general television/film components like on-screen movement, image fidelity, and scene-changes. These discussions lead naturally to an investigation of film-techniques in computer interface, as a possible tool for increasing or controlling the impact of agent-driven messages.

Studies on Faces

Several researchers have studied the affects of anthropomorphic representations (and, specifically, human faces) on user comfort, perceptions and work-habits. Walker found that giving the interface a face produced a more engaging effect, but also required more attention and effort from the user. She also raised the idea of using markedly different faces for different tasks, such as providing information and requesting information. [Walker94]. King explored how people react to shapes as well as faces. He found that people attributed the most intelligence and agency to 3D human faces with blinking eyes, but that various representations, rendering styles and motion also played a role. Also inherent in his work is the caveat that the representation should match the abilities of the agent – if a photorealistic face is used, it implies that the agent has human-level intelligence and communication skills. [King96][Koda96]. Takeuchi compared an abstract symbol and a synthesized face as the means of providing helpful information. Within the context of a card-game, he found that the face was more engaging and fun, but required more attention on the part of users. The more abstract object (an arrow) allowed users to focus on the card-game itself. [Takeuchi95].

Computers As Theatre [Laurel91]

Brenda Laurel has recognized the usefulness of theatrical and artistic input to agent technology for many years. Her work approaches human-computer interaction from a basis in communication we've built over centuries of experience, drawing parallels to stage and set design, character representation, and theories of dramatic storytelling.

In discussing anthropomorphism, she points out the usefulness of character in communicating predictability and the display of intentionality – one of the more difficult questions arising from any discussion of user-comfort with autonomous software agents. Indeed, "in most cases, we want to be able to predict [agent] actions with greater certainty than those of 'real' people", an idea which falls well within the skills of the acting profession. [Laurel90]. She also notes that gestural-language work, currently studied as input [Brennan94][Buxton90], has potential as output as well. Verbal natural-language features can be augmented or supplanted by the conventions of body language we use in daily life as another form of "natural language". Embodied characters, than, enlarge the vocabulary with which we can communicate with the user.

Her work touches also brings up other relevant ideas: by approaching the problem of computer interaction with a dramatic bent, she opens up the possibility of a direct application of cinematic and dramatic techniques to interface display and design. The techniques of a set-designer can be brought into play, to communicate mood, tone, or tension much as they do in a dramatic work. Rather than simply taking graphic screen design a level further, however, she considers the very action of working with a computer a dramatic one, with communication occurring at a variety of dramatic levels and meanings.

Film Craft in Interface Design

Chuck Clanton and Emilie Young analyzed the overlap between computer interface and film techniques, with an eye towards incorporating the emotional and communicative impact of these narrative media into traditional computing. They addressed ideas such as:

-Framing: activity at the centre of a frame has different meaning than action at or directed towards the edge of the screen. Placement screen-right and -left also implies entering/exiting or moving forwards/backwards in time. Adding dynamism to the way multiple object are displayed, such as a camera shot of one character re-adjusting to a two-shot when another enters, can be used to suggest a relation between the objects.

-Camera Techniques: the distance from a camera to an object can show importance, as can focus. A sequence of close-ups can be used to show a causal relation between two objects, and draws immediate attention to them. The angle of viewing also affects how the viewer feels about the object (ie: low-angle shots indicating power and dominance, while an overhead view signifies remoteness)

The techniques discussed draw tend to draw parallels to existing projects rather than suggesting new techniques or directions for research, but remain an excellent example of how closely allied the two fields may in fact be.

chapter three

Theories of Animist Interface

Over the course of studying and experimenting with animist interface agents, certain ideas emerged as useful for enriching agent-user interaction. What follows is a discussion of the decision for Animism (as opposed to anthropomorphism), the importance of foreground/background elements, the graceful degradation of communication, lighting/camera work, and sound.

The Case for Animism

If one is to believe critical literature, Anthropomorphism and Agent-Embodiment seem to go hand in hand. Tirades claim that agent-based software emasculates the user by acting as a "person" and thus relegates the user to the role of a computer [Lanier95], and that treating machines like people will make us treat other people like machines [Schneiderman96]. Beyond suggesting difficulties with agents in general, these arguments implicitly accept a humanoid representation (or at least one with human-level capabilities, such as Peedy the talking parrot [Ball96] or Microsoft Bob's saccharine fuzzy-animals.) as the inevitable outcome of any attempt to visually represent the agent-intelligence.

This idea seems inherent in fictional, conceptual and Artificial Intelligence approaches to agents. In his 1985 American Association of Artificial Intelligence Presidential Address [Bledsoe86] Woodey Bledsoe told of his continuing dream to build a computer friend. He spoke of the "excitement of seeing a machine act like a human being, at least in many ways," of building a machine that could "understand, act autonomously, think, learn, enjoy, hate" and which "likes to walk and play Ping-Pong, especially with me" [Bates94].

However, this idea has produced many of the more marked failures in personified interface. For example, "bank terminals have matured through the early phase in which anthropomorphic designs (Tillie the Teller, Harvey Wallbanker, BOB the Bank of Baltimore) dominated. The anthropomorphic styles are cute the first time, silly the second time, and an annoying distraction the third time." [Schneiderman96]

Animism is the finding of life (or "animate behaviour") everywhere around – the attribution of consciousness to nature and natural objects. Encyclopaedia Britannica Online defines animism as a "belief in spiritual beings who are concerned with human affairs and capable of intervening in them" [Britannica97]. Experiments in interactive characters have produced many enjoyable, coherent and believable characters without a human or caricatured representation. These "lower order" lightweight forms, from animals to plant-life, or even natural qualities such as light, seem to be interesting areas for interface experimentation. The value of the agent-style thinking remains untouched: literature making the case for agent-based thinking notes that "The [traditional] rote-instruction follower image affects and infects our thinking and limits our imagination about what computers can do." [Travers96], and that

> "Animacy, then, is more properly understood as a framework or way of thinking…Animate thinking stems from a basic need to explain happenings and tell simple stories about them, and a need to fit things into roles in the stories as actors and objects of action" [Keller85]

None of these claims imply a need for natural-language output, or complex human-like facial expressions, or any of the other accoutrements of anthropomorphic representation. Ibister notes that "assessing whether a computer agent is 'intelligent' is not a matter of deciding whether it has human intelligence or mimics human neurological functioning. Rather, it is a matter of assessing whether the agent possesses the cues that lead a person to label someone/thing as 'intelligent'." [Ibister95]. She goes on to suggest using other cues to create human-like characters, but the point is still well-taken when applied to non-human characters. Indeed, lower-order life forms may prevent many of the over-attribution errors a human-like character inspires. Making a character speak in full sentences implies that it is capable of understanding the same – in most cases, an incorrect implication. Similarly, support for human-like reactions and photorealistic imagery implies the system can interpret and display sophisticated human emotions. Laurel warns that agents should have "a symmetry in modalities" [Laurel91], matching abilities to appearance.

If, on the other hand, we are presented with a mapping between the health of a digital plant and, say, the health of hard-drive, it doesn't imply that the plant understands what or why a problem is happening. The plant can display state without creating false expectations about how the user should deal with that state. In terms of predictable and characteristic activities, however, the lightweight characters are much more tenable. A dog, for example, is much easier to fit within a well-defined, well-understood stereotype:

"The idea of a canine agent…suggests a class of activities (fetching the morning paper and announcing intruders, for instance), a level of competence (the agent will deliver the paper but will not provide commentary), and a kind of communication (a small repertoire of simple commands and an equally small repertoire of behavioural responses) that may be entirely appropriate." [Laurel90]

Essentially: we don't have any familiarity with parrots that can work Jukeboxes, or cartoon heads that can read our email. We do have experience dealing with pets, or accepting lighting conditions and setting as indicators of emotional atmosphere. These "lightweight" characters can also be worked into multi-agent systems with greater ease. Different aspects of system intelligence and communication can be reflected by different embodiments. The other two options, creating one comprehensive super-interface-agent or having multiple competing fully-rendered characters, seem much more confusing. Artificial characters from the entertainment world seem well developed for engaging and intriguing users – it seems useful for us to incorporate the techniques that work when designing interface agents.

Foreground/Background

The Ambience project of Hiroshi Ishii's Tangible Media Group at the Media Lab has been the source of much inspiration and guidance. His team focuses in part on background media -- information conveyed through subtleties of sound, movement, and lighting in the physical world around us. The ideas also have merit inside the world of the computer: agents don't always need to communicate as if they were the centre of attention, but can instead use imagery, colours, screen-lighting or sound to communicate through background channels. The user can remain focussed on their tasks, yet be kept appraised of whatever the agent may be doing (handling network security, translating/parsing mailing lists, etc)

Indeed, some of the problems voiced by critics of agent-technology may be answered by these ideas. The user can be made more aware of what the agent is doing on their behalf, in way that doesn't interfere with their own tasks. This style of communication is not very common in today's software; system diagnostic tools exist, for example, but the user has to query them to find any information. Intelligent software can poll the very same diagnostics, and use non-intrusive techniques to convey this state. The feelings of being in control, or at least well informed (rather than being alerted only when the agent needs approval) can help the user feel more comfortable with the automation and autonomy inherent in agent-style software.

This fits within the idea of an "animist interface". The state of the software and the system is embodied, but not as a humanoid character or free-floating face. The very environment takes on a meaningful function. Presently, many of the visual elements on the desktop are set arbitrarily – at best, they are used as a way to personalize a computer. Beyond this, the colors and imagery remain completely static, despite their visual prominence and ubiquitous presence. By changing desktop-patterns, window borders and color-schemes from arbitrary bitmaps to useful interface-elements, we can inform without adding to the visual complexity of the average computer-desktop.

Furthermore, these background or "ambient" communication channels can do what few character- or facial-based systems can: communicate effectively without needing the user's full attention. "Most of the visual field is in peripheral vision, and peripheral vision has a special deficiency – image fidelity. When vision is fixed on any particular word, facial feature, or object, most of what we see is rendered quite poorly" [Reeves96]. With this in mind, it seems limiting to tie all agent-driven messages to a detailed face. Indeed any complex rendered object seems problematic and distracting. If we can deliver general messages about mood or intensity, danger or change with low-resolution, subtle and environmental changes, we can increase information flow without overwhelming the user. In the same way that we pick up on subtle cues in our real-world surrounds, or implied messages in the mise-en-scene of a film, we can receive information through the background channels of our computer screens.

Within this environment, messages can still push into the foreground: characters or elements can suddenly shift opacity, coloring and location to attract the user's attention as deemed "necessary" by the agent-intelligence. However, we can again use "background" channels to augment the information being delivered. Rather than using traditional full-text dialog boxes for every message, no matter where it originates or what it says, we can use the contents of the message to affect how it is delivered. At our disposal are the tools of computer-based graphical design: color, rendering style, visual prominence, animation, etc. Jagged, "nervous" fonts and imagery might indicate a security problem; a character might tell you about a mail attachment while sitting near the application needed to view it; messages from distant locations might be reported by an exhausted, drooping figure; lighting conditions might dramatically shift when the user is alerted to a possible system-crash.

The message conveys less information than a full-text dialog box might. However, this is a purposeful ambiguity: the message is quickly digested when its not immediately relevant, so false-positives are not too distracting. Furthermore, it can display information from a variety of sources without using one explicit gauge/indicator for each feature. When necessary, the user can act to get more information or the agents can push into the foreground. The combination of these techniques can blend agent messages more seamlessly and usefully into the user's workspace.

Graceful Degradation

Well-designed, beautiful, multi-media interfaces must walk a delicate line between style and utility. The emotional impact of using the system must mesh with the actual functionality at all times. The problem, then, is that characters and settings should maintain the impression of coherency and consistency while still performing useful functions.

In order to preserve an illusion of life, for example, Microsoft Agent's Genie runs through several idle-state animations. When not otherwise instructed to perform some action, the character will nod off to sleep, or pull out a pen and write something down. Unfortunately, this leads to immediate interpretive difficulties: when is the character communicating useful information, and when is it just running a consistent animation? One viewer of a system developed during this thesis asked "Why is he bored? Am I boring him?". The answer is straight out of the Velveteen Rabbit: "No, he's just wants to be Real." [Williams84]

Highly anthropomorphic characters enhance communication through animating facial expressions, body language, etc. However, as a result, they have too many degrees of freedom to occupy when they have little or nothing to communicate. An interface agent is not necessarily the same as an AI-complete lifelike character, nor does it have the complex inner motivation and emotions the way an actor in a film might.

In order to remain more than a distraction, the character needs to degrade its presence -- to fade into the background, out of notice, in a consistent and graceful manner. The character should use dramatic movements only when it has something to say; it can also use a steady animation which continuously displays state. Simply popping in and out of the foreground like a dialog box, however, restricts the system from conveying a range of information. Thus, control over opacity, size, colour saturation, location, graphic detail and movement can all be orchestrated to fade the character in and out of the foreground as what it has to say becomes more and less relevant. Essentially, we want to be able to scale presence to match importance.

To this end, we must also choose our characters correctly. If fully-rendered humanoids have too many degrees of freedom for the agent-intelligence to animate meaningfully, then perhaps a lower lifeform or some other recognizable and readable an embodiment is called for. Plant life, for example, can have a fairly obvious set of mappings between health (or form or colour) and "mood". Traditional animators have produced wonderful and communicative characters from archetypes as simple as a Persian rug or a flour-sack. With this kind of character, we don't control actions, but emotions. The simplistic nature of the representation controls what it does, and outside stimuli influence how it does it.

A lightweight character may have a more natural idle-state, or more obvious characteristic activities it can perform to preserve the illusion of animacy. The simplest, most readable character should be chosen for the task at hand.

"Lights, Camera...Action!"

In the movies, actors are not dropped in front of a camera and told to act. The scene is not played out in whatever location is handy, with a couple of halogen lamps plugged in near-by. Rather, the look and feel of a film derives from the confluence of imaginative and skillfully-executed camera-work, costuming, lighting, set design, and a host of other skilled talents whose credited perpetrators rush by in 12-point font while the closing theme plays. These areas suggest incredible opportunities by which the interface designer can increase the impact and legibility of their characters and settings. Furthermore, they can serve to increase the explicit nature of a message; as Hoffner and Cantor noted in their study of the perception and responses to media, "formal aspects of television and film productions may enhance the salience of certain nonverbal behaviours."[Hoffner91] The language of drama is at times the language of melodrama: an enhanced, stereotyped vision of emotion and communication painted in broad, readily apparent and easy-to-understand strokes.

Movement of a virtual camera, for example, seems to be an idea completely ignored in existing character interfaces. Proponents of anthropomorphic design often make the point that we as people are used to the dynamics of social interaction, and therefore could benefit from computers that responded to and replied with the same language. The "language" of cinema is also extremely familiar to us, and has created a very sophisticated set of techniques for implying information to the audience. Cuts, pans, fades, L-cuts, focus-shifts, framing, and the wealth of other techniques the expert cinematographer have at their disposal can be imported wholesale into the world of character interfaces.

If a character has an important message, it doesn't need to jump up and down and wave frantically; it doesn't even need arms or legs as part of its representation. A well-timed cut to close-up, or a sudden zoom, can express more about what that character is "feeling" or what message it is trying to convey than a dialogue or text ever could. Framing in that close-up could establish a relationship between a character and an icon -- bringing a cinematographer's eye to the layout of the desktop as well as the movement of a character. A series of fades could be used to signify time passing as a character is involved in a long download, or a slow change in focus and depth of field might indicate progress in a long file-copy. A series of fast cuts, or a hand-held effect in camera-movements could imply high-tension situations like insecure data-transmissions or automatic running of anti-virus software.

Lighting can be used as an ambient communication device, for example showing the passing of time by movement of a light-source. However, cinematic techniques of lighting design can also be used to add dramatic effect to communication. A mobile character can be lit from beneath to increase tension when asking for confirmation of important instructions, or backlit to heighten mystery during an untrusted applet-download. Changing the colour and angle of a virtual lighting source could emphasize that a character was switching topics or contexts.

For example, both Sheth's "Newt" news-filter and Kozierok's calendar agent used simple faces/characters to build trust – they kept the user informed about agent recommendations by making it clear when the system wasn't very sure about what it was producing [Sheth94][Kozierok93]. Both of these projects might be enhanced by techniques of lighting and camera-moves designed to present the character in whatever "light" is required.
In another example, Microsoft's "Office Assistant" Paper clip lurches forward and taps at the window/"camera" when it urgently needs the user's attention. The framing of the

character emphasizes emotion, and the z-axis movement attracts more attention than 2-D animation would. This effect is Assistant's more useful animation, as it performs an action none of the other icons in the interface can duplicate [Microsoft97].

Essentially, the idea is that a computer-generated "world" is *not* bound by the same rules as the real world. Time, location and movement are malleable, and we can use existing and well-understood cinematic techniques for telling stories or conveying information if we recognize this. The setting can be used to increase the bandwidth of communication between the computer and the human. If the point of embodying computer state is to make it more understandable, then we should use whatever languages, techniques and tools are available to get our messages across.

Sound and Music

The use of sound is an important and useful part of building a coherent and understandable space. In tests on perception of "emotion" in a computer/video character, Borkenau found the quality of audio to be more important than the quality of video images. He also noted that specifics about the audio information, such as 'unpleasant voices' or 'halting speech' were noted much more than visual cues [Borkenau93]. Incidental music and foley sound-effects can convey information through a channel that doesn't distract the user from the visual task they're engaged in. Today's midi-capable personal computers still limit the audio-component of their interface to a single-tone "error" beep, or some slightly more complicated instantiation of the same functionality.

Beyond decisions to avoid technology such as speech synthesis (the current quality makes it more distracting then useful -- research suggests that low-quality audio is more distracting and noticeable than low-resolution imagery [Reeves96]) -- time constraints dictated that sound not be more fully explored in this thesis.

These theories were developed while brainstorming and experimenting with many related ideas. The next section details the three experiments that produced and incorporated these principles.

chapter four

Three Experiments

The ideas of Animist Interface were produced through a design/implementation/iteration model of development. After researching relevant areas, I sketched and implemented an idea. I analyzed that project, found interesting points, and re-did the entire process. That second idea was re-analyzed, and I came up with a final project that incorporated salient points from other two. In all cases, the "functional sketches" were developed to point of usability, not full functionality. I produced just enough intelligence to provide interesting output for the display algorithms.

The next few sections describe the three prototypes, addressing issues of design, implementation, and critical review. "The Genie", an anthropomorphic interface, was developed to explore the issues of human-like characters on the computer desktop. An analysis of the dangers of that style of interaction led me to experiment with the "Background" project, replacing the desktop background with controllable random-access full-frame video. Finally, "Mise-en-Scene" experimented with lightweight non-anthropomorphic characters in a rendered environment.

Experiment 1: The Genie

Overview

(Note: some this section is adapted from [Lachman97])

Most of the attention surrounding embodied agents is concerned with "traditional" character-based systems; that is, with recognizable animal or humanoid forms (albeit of a variety of rendering styles and animation types). However, flagship systems such as Eager [Cypher91] Newt [Sheth94] seemed to use the character as mere indicators of static state: the system is listening, or it's busy, or it has a suggestion. Even the animations of latest-generation systems like Microsoft's Office Assistant [Microsoft97] are focussed on simplistic communication like "I'm now printing" or "I realized that you told me to save the file, and I'm now saving that file". They convey that information in an entertaining, well-animated fashion, but the goals they are pursuing are limited and not that helpful. As Disney animators caution, "Do not write dialogue that describes what you are seeing. If a character is panting and sweating form the heat, the last thing he needs is a line that states, 'Boy, am I hot'" [Disney p387]. A character's animation should not be designed to convey what the user already knows.

If one has taken the admittedly risky step of incorporating a character-representation into software, that representation should add fundamentally significant features to the system. Feedback about the machine correctly receiving and interpreting commands is indeed a useful feature; however, it doesn't justify a major metaphoric shift in representation. Animating the Office Assistant in order to get your attention (ie: the Paper Clip character tapping the screen and peering out at the user) seems closer to the mark, but still doesn't capitalize on a key concept. What an anthropomorphic or animalistic interface adds to the equation is the ability to *act*. To perform. To convey a message with emotion, with body language, with meta-information that lets us know something about the situation even before we consciously parse the message. There are many ways an animate creature can display this meta-information, such as representation, rendering, or style of movement. Each method can communicate while implying information about what the message is. What character-based interfaces can leverage is our ability to read such meta-information transparently, at a sub-conscious level.

We can also use the setting, the environment in which the character interacts, as part of the communication process. This can be expressed through animated and expressive environments, or by having the character react to the computer-desktop – in effect, placing the user and character in the same environment. Gaze direction, location and other features of the character's animation become more meaningful when there are other features for the character to "play off of".

"The Genie" was the first design-sketch produced. My goals were to experiment with a decidedly anthropomorphic interface that was capable of several responses to a particular message-type. As part of the project, I also wanted to create an intelligent agent back-end to drive the process.

The system directs a cartoon-rendered Genie on a Windows desktop, which performs animations at the request of a web browsing and an email-parsing agent. The character's behaviour is based on a rating of how hard it should work to attract the user's attention, not just on the fact that it has a message. The character is a frameless window, controlled through instructions a server sends to a web-browser window.

In the first example, email received to an account triggers a sender-lookup (see *Email Filter* in the Implementation section for a discussion). The priority-level associated with the sender is used to influence the character as it attracts the user's attention to their desktop email-client (if it chooses to do so at all).

In the second example, the character is connected to Maitre-D, a site-based web-browsing agent built as part of my agent-based research. The user can browse a specific location with Internet Explorer -- in this case, the 819 pages of the Interactive Cinema site. Their actions are reported back to software on the server, which comes up with a rated recommendation of other pages on that particular site which the user might like to visit. Instructions for character-behaviour are sent as JavaScript back to the client, and the Genie reacts accordingly. The user can click directly on the character to see and act on Maitre-D's recommendations.

Implementation

The system uses the beta-1 release Microsoft Agent ™ Active-X control for animation and character appearance. The Active-X control is a windowless animated Genie, which can be controlled through a C++, Visual Basic, VBScript or JavaScript interface. Instructions include such commands as "GlanceLeft", "MoveTo(x,y)", "Approval", and "Sad". In this project, the genie is controlled through JavaScript dynamically generated by Perl 5.0 CGI and shell scripts, served by a Netscape Communications web-server.

An instance of the Internet Explorer web-browser initializes and controls the Genie, holding it at a neutral state and sending it to an initial location. The window acts as a conduit for all information to and from the actual agent-based intelligence, which for our example resides server-side (although in practice the intelligence behind the character could be realized using a combination of local and remote influences). The browser is constantly polling the server for instructions, and reporting the state of the client -- for example, if the user uses another browser-window, the browser reports when a new URL is loaded.

The server returns instructions in the form of new JavaScript function-calls, either maintaining the Genie in a fixed state, or instructing it to move and express an action. Various server-side subsystems are polled to determine the commands (eg: an email-filter or the Maitre-D site agent [discussed below]), and actions such as web-browser commands, outside the scope of the character's animation, can also be returned.

Note that the Agent control has several features, such as voice command and response, and idle animations-loops, which have been disabled for the purposes of these systems. For a discussion of the reasons behind the configuration, see below.

Email Filter

The Genie can react to email sent to a certain account, acting as an intelligent "xbiff" (a program used on Unix systems to alert the user of new mail). A flat-text file was written, containing email addresses and associated "priority" ratings. High-volume lists, for example, were given low-priority, while close friends or the author's professor received higher rankings. The list simulated the work of an intelligent user-modeling agent, which might be able to glean such rankings through observing user-actions over a period of time, or by communicating with other user-agents in the same workgroups and net-communities [Maes95].

When email is received at the user's Post-Office server, it is stored for delivery as usual. A Perl script scans for the highest-ranked incoming message in the queue, determines behaviours according to a list of reactions, and generates JavaScript instructions. These instructions are sent via HTTP to the Active-X Genie on the user's machine, which animates accordingly.

The reactions include such behaviours as simply glancing at the icon of the user's email-reader for low priority messages, or pointing more vigorously for higher rated email. Top priority messages cause the character to wave its arms to attract attention, move across the desktop to a position next to the email-reader application, and stare at the icon (Figure 4).

This example places the character "in the world" of the computer screen, allowing it to interact with the same icons, applications and desktop tools as the user. The idea is to allow the character and the user to use the same vocabulary and functionality, to encourage the impression that they are both approaching the same problem set (although with different skills) – in effect, to establish the common "grounding" Susan Brennan notes as incredibly important in our human-human communications [Brennan90].

This is a departure from systems which try to make the specific application inseparable from the interface-agent, as in the case of Newt [Sheth94] or Apple Guides [Laurel91]. Nass and Reeves have found that presenting a character (or the computer itself) as co-equal, a member of the same team, increased the user's trust and intuitive understanding of the agent and its actions. These kinds of feelings are extremely important to instill in someone using a software agent, and at present are less well understood than, say, knowledge-acquisition, or user-modeling, or other intelligence-based aspects of agent creation. This sketch represents an interesting direction for dealing with comfort-levels in software agents.

By freeing the character from a particular location on the screen, mobility becomes part of the Genie's expressive vocabulary, which in turn extends the system's ability to communicate. The character has more maneuvers for gaining the user's attention in its repertoire, and more ways for the to system communicate about specific applications. As users and machines multi-task, the interface-character needs to switch contexts just as rapidly: referring directly to the screen elements and icons that the user is manipulating helps communication to be efficient and clear.

Maitre-D Interface

Maitre-D is a web-site based browsing agent being developed by the author. Through off-line textual analysis of the pages associated with a site, Matire-D creates cross-indexed relations between similar pages. As a visitor browses the site, Maitre-D notices which page the user is on and is able to produce rated recommendations of related URLs the user might want to check out, all on the same server. The proposal is to make visiting a web-site into a two-party agent-interaction: a user-agent, with a model of the tastes and current interests of its user, negotiates with the site-based agent and its intimate knowledge of the content it hosts. The result would be a rated recommendation (or group of recommendations, as a configurable option) which can then be passed on to the user.

An animated character can work in concert with Maitre-D, as the interface to the intelligence. The Active-X character is spawned by a second browser-window, which also contains embedded JavaScript functions to track events in the user-controlled window. Whenever the user moves to a new page, the new URL is reported back to Maitre-D, which runs server-side. Browser history, keywords which describe the new page, and two "reserved" fields are then used to produce a group of recommended "next pages" for the user to browse. The two extra fields simulate the input of a personal user-modeling agent: one contains current keywords that make up the user's general profile, the other short-term interests for this particular session. A user-agent could fill in these fields with a variety of subjects, depending on both how much it wanted to reveal about its user and its perception of the user's intentions during the site-visit. The strength of Maitre-D's highest recommendation, based on these inputs, is used to create behaviours in the client-side character.

For example, low-rated recommendations produce no visible reaction in the character. If the user wishes to view the agent's suggested pages, they can click on the character directly for a plain text, ordered listing. Better matches can cause the Genie to quietly nod once, to look around and shrug, or to wave his arms to attract attention (Figure 5), depending on the numerical ranking of the recommendation. Again, clicking produces a list of recommendations, and selecting from the list sends the browser to the new page

Figure 4: Genie with Email Agent

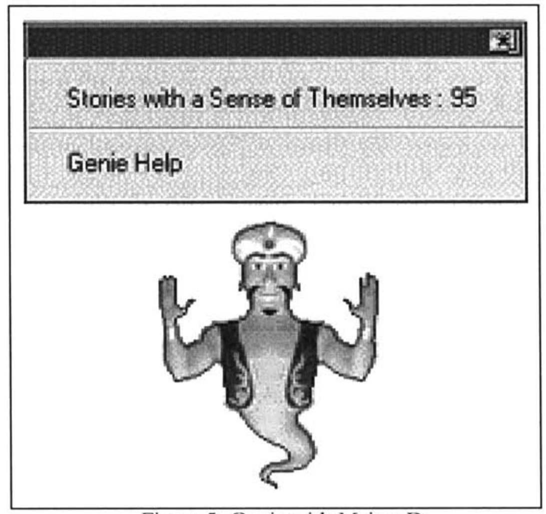

Figure 5: Genie with Maitre-D

Analysis

There are several difficulties with the email system as it stands. First of all, the agent communicates through simple body language. This can certainly be a problem in a "world" with multiple stimuli -- gaze-direction is a fairly arbitrary reaction to the specific domain of email handling. However, as the perceived importance of the message increases, the ambiguity of the character's response decreases. The character shifts from subtle body language to gross locomotion; communication occurs along a spectrum, with gestural communication at one end and the direct stimulation of the human response to dramatic movements on the other.

When connected to Maitre-D, the character has a decent range in its behaviours. Vigorous gesticulations draw immediate user-attention, while subtle movements are visible only if the user is paying attention to the Genie. The discrete nature of the communication, however, remains a problem. If the initial movement is not seen, it isn't repeated. Rather than influencing the mood or general behaviour of the character, the control can only apply stimuli to distinct, discrete animations. Communication on the periphery of attention remains unexplored.

This particular project produced several recommendations for anthropomorphic interface designers in general (and the Microsoft Agent team in particular). First and foremost is the point made at the beginning of this section: the primary benefit of an embodied interface lies in the range of possible expressions. Characters from full 3-D figures to line drawings can all produce a wide variety of emotional poses; discrete, frame-based animation completely ignores this fact. In order to capitalize on moods, to add in parameters for how to do something (in addition to "what to do"), other animations styles seem called for. Spline and NURBS graphical styles will be discussed in a later section [see "Mise-en-scene" below]

Related to this is the need for an architecture that supports multiple directives. As [Blumberg96], [Elliot95], [Elliot97], [Perlin96] and others have discussed, some combination of action-skills and emotional modifiers can increase the range of acting abilities of the character. The ability to "GlanceLeft" **and** look "Sad" is a useful, if not essential, trait. A greater granularity of control in the character results in a more sophisticated control-interface; as already discussed, however, it seems hard to justify the presence of the character at all without it.

Finally, the messaging system should be designed around a range of priorities and types of messages. If all vocal messages in the real world were delivered at the same pitch, volume, and timbre, it would be impossible to glean the meta-information we need in order to make sense of the world. In the same way, the palette for graphical characters should be fully exploited to make the most of the medium: specifically, features such as scale, perspective, opacity, and even graphical style can all be used to add to the range of messages the character can deliver. The Microsoft control is a great prototyping tool, primarily due to the quality of rendering and ease of programming. However, its underlying architecture restricts efforts, implying that most characters created with the system will look and behave almost identically.

This project began with the premise that anthropomorphism was the form of expressive-interface most suited to agent intelligence. From here, I realized that many of the benefits I noted could be embedded in other, less "dangerous" representations. The Genie itself seemed too striking a character –too complex and too detailed to convey the relatively simple messages produced by my software agents. And yet, animated icons (like the mailbox-and-flag indicator found on some Unix systems) were too special-purpose. Such icons are basically just gauges, taking up space when they don't have much to "say", and communicating with a very limited range of expression when they do. It was only after this analysis that I began seriously working on characters with non-humanoid appearance and abilities. I wanted "lightweight" characters whose activity and presence could "degrade" more readily; however, I wanted to maintain the expressivity that non-iconic representations gave me. Animist characters should still be able to move across the screen, and interact with icons (again, setting themselves up as teammates, working in the same "world" as the user). Many examples of character and setting could work in concert (ie: treating the lighting and setting as supporting-roles at some junctures, as centre-stage at others) to provide both foviated and peripheral communication. These ideas led directly to the next experiment, "Backgrounds", which in some respects lies at the opposite end of the "character" spectrum.

Experiment 2: Backgrounds

Overview

Colour, sound, and rich graphical/iconic representations characterize today's computer desktops, replacing the system-font, monochrome displays of the past. But as Shneiderman notes, "the computer industry appears to have inadvertently created a de facto standard based on 1984 era hardware and window managers" [Shneiderman97]. Certainly, background imagery or "Microsoft Plus!" sound-and-graphic schemes reflect personality and customization, and can make the computing environment more inviting. The potential for using these elements as an integral part of the computing interface, however, seems relatively unexplored.

The second sketch deals with harnessing these "background" graphic elements as part of an interface, to convey continuous information through subtle means. Processor load, network traffic, general information about system-health, or time-variant information like laptop battery-life can be interpreted through the movement and content of desktop-background; download completion or background processing is reflected through existing interfacedecorations such as window-borders. For this experiment, the processing is done on a separate networked machine, to prevent the interface from affecting the very system it's trying to characterize.

Implementation

The VGA screen-output of a Pentium machine is passed through a signal-converter, and sent as video to an SGI Indigo2. A Galileo video-processing board allows the SGI to chromakey out the Windows NT background (configured to a solid colour), and allows software on the SGI to blend in arbitrary graphics and video.

The Performance Monitor package on the NT allows the state of various hardware and networking devices to be polled; it passes along information such as how hard the processor is working, how many foreign users are accessing local drives, or how many packets the network-layer is dropping. A Visual Basic program writes this information to a network drive, shared with the SGI under a PC/Unix file-sharing system called Samba.

The SGI controls the clip selection, playback speed and play-direction of various video-files, based on the information about how the PC is behaving. The NT desktop and the video are blended together and output to a video screen, which is used instead of the regular PC monitor. The clips include time-lapsed imagery such as city-scenes (automobile traffic, the sun rising and setting on a crowded street, etc), cloud movements and the apparent movement of stars over the course of the night (Fig 5).

In addition to replacing the desktop background, the chromakey blending itself can be tweaked, allowing the video to be seen as subtle modulations in application-windows, or at various opacities while on the desktop itself.

As a small additional feature, simple controls were written (again in Visual Basic) to modify system-properties such as the thickness of window-borders and title-bars, as an alternate (and sometimes less obtrusive) means of information display.

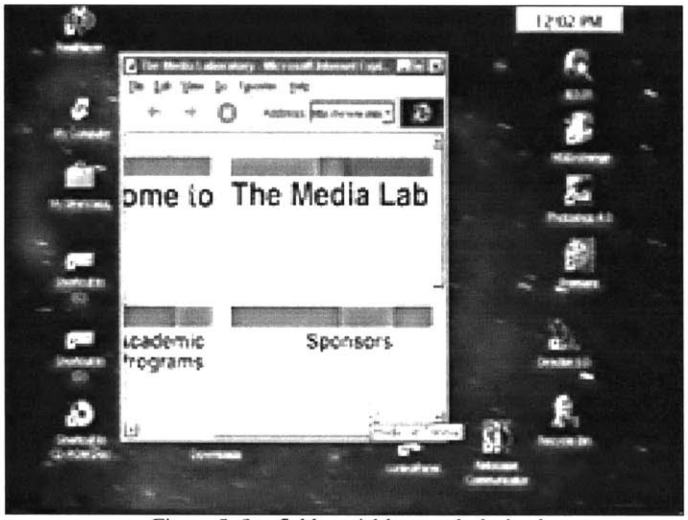

Figure 5: Starfield, variable-speed playback

Analysis

The system was designed to harness "dead" space in the interface, conveying continuous information without taking up screen real estate. Rather than special-purpose diagnostics, requiring the user to *decide* to examine system performance, static elements of the screen could be incorporated into the interface. Communication doesn't have to happen through new elements popping into an x-y location or onto the window-stack, but can instead be incorporated into the features already visible, for a more efficient, smooth, coherent interface.

However, there is a deeper idea involved, which reaches beyond an intelligent use of visual space: a subtle projection of continuous information, keeping the user informed of the state of a system through background channels. The kind of information measured by this system is really only "important" at extremes -- when the processor is working hard, or when the network is about to hang. The rest of the time, a general impression of how things are going allows the user to troubleshoot or take pre-emptive actions (ie: "The filesystem isn't doing too well -- its not a big problem, but I might as well take care of it while I have spare time"). The "display" of information has low resolution, granularity and specificity, but a very high visibility and affect.

Existing gauges and diagnostics programs are examples of displays that don't have any dramatic representation of dramatic information -- they are either run at distinct intervals, when the user decides to check on performance, or with binary communication, throwing up dialog boxes when some discrete error-state is reached. In contrast, environmental changes are designed to display information continuously, becoming more visible as the information they convey becomes more relevant to the user. For example, the video-playback becomes jittery and fast when the processor is working extremely hard; the lighting of a city-scene dims rapidly when the network is lagging. The system state, then, is always available, but becomes more apparent as it becomes more important. The setting of the computer screen -- lighting, imagery, style -- can be used the way it is in film: to subtly convey impressions, feelings and moods without distracting from the main action.

The problem, of course, is that the background can be more distracting than useful. The imagery I selected creates movement over a large area of visible space, with an almost hypnotic effect; the backgrounds need to be more controlled, and to degrade to a low-level of activity more gracefully. Again, the goal is that low-priority information should have low visual impact, which the full-screen video doesn't completely meet. Greater control over what the imagery is, as well as how it reacts, might help solve this problem.

Figure 6: Background with high processor load

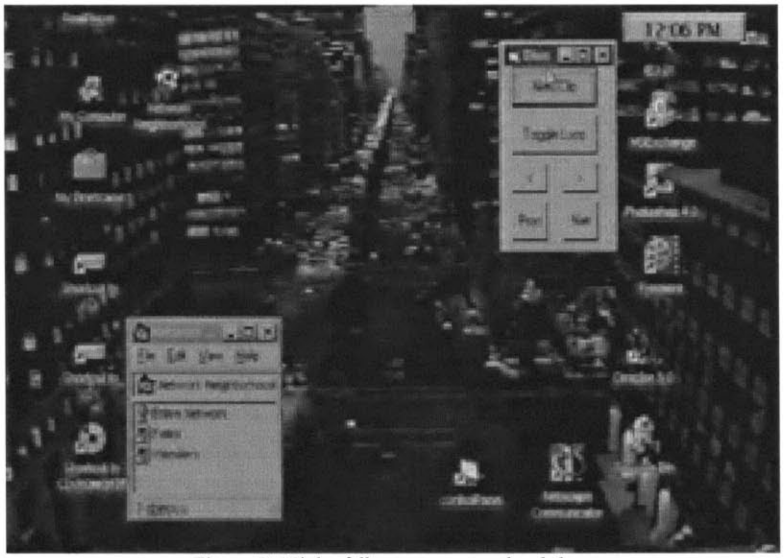

Figure 7: Night falls as processor load decreases

The nature of these background images is also very important. At one time additional clips included very "loaded" imagery, such as a bonfire and a kabuki-theatre actor. Some clips had far too many associations to blend with the interface well, given that the current system pays no attention to clip-content (ie: showing processor-state with fire is a very strong statement; if the agent doesn't have the intelligence to use such statements wisely, it really is misleading the user). The right course seems to be one of the following:

1. Select abstract, more "neutral" imagery;
2. Only modulate existing backgrounds. Let the user continue to customize the content according to their personality or whim; or,
3. Build the intelligence to correctly use a database of "tagged" clips and images. Select appropriate content to display as part of the message.

In addition to using these ambient techniques which communicate in their own right, I wanted to experiment with using them in a support role. "Mise-en-Scene", the final experiment, was designed to explore the realm of character/agent in a dynamic environment.

Experiment 3: Mise-en-scene

Overview

The environment around an actor may be just as relevant to our understanding of their thoughts, intentions or feelings as any "bit of business" the character performs "center stage". A crowded city-street with a lone, still character implies a certain emotion, as does the pathetic fallacy of a grey rainfall and a heartbroken lover. The movie-camera doesn't merely film characters wherever they happen to be, but records a complex interplay between actors, scenery, lighting, and the framing and production-techniques of the final screen image. Quite literally, the theatrical term "mise-en-scene" refers to everything which is "put in the scene" – actors and lighting, the physical setting, and their positions within it.

The third example sketch attempts to meld the idea of a simpler, less-anthroporphized character with the embodiment of information as environmental condition. Lighting, rendering styles, movement, camera positioning and framing are all used to build moods, which can imply frantic activity, risky security conditions, poor system health, etc. More direct messages can cause the characters to take on more anthropomorphic tendencies as needed.

Implementation

The chromakey/blending setup from the previous sketch is used for this project as well. The characters and environment are rendered Inventor 3-D graphics running on the SGI, again blended into the desktop of a Windows NT machine and output to a video monitor.

A NURBS (Non-Uniform Rational Besier Spline) [Werneke94] surface and curve are used to make two simple "characters" (Figure 8). Unlike frame-animation, in which all variations of the animation must be pre-rendered, or polygon models, in which geometric primitives are manipulated through affine transformations, the NURBS constructs allow for real-time control over many details of rendering and positioning. The control-points for each character can be manipulated to animate and move the character, but can also, say, make the curves more spiky, or smoothly transition into a completely different shape (Figure 9). The very nature of the surface can be animated, rather than simply moving separate limbs – the idea being to increase the expressiveness of the entire rendering without having to produce time-based discrete actions which the user must notice as they happen. When the characters do have the user's attention, we benefit from a reaction that Gonzalez's study noted: smooth animations allow users to make more accurate judgements about what's happening than instantaneous transitions [Gonzlez96].

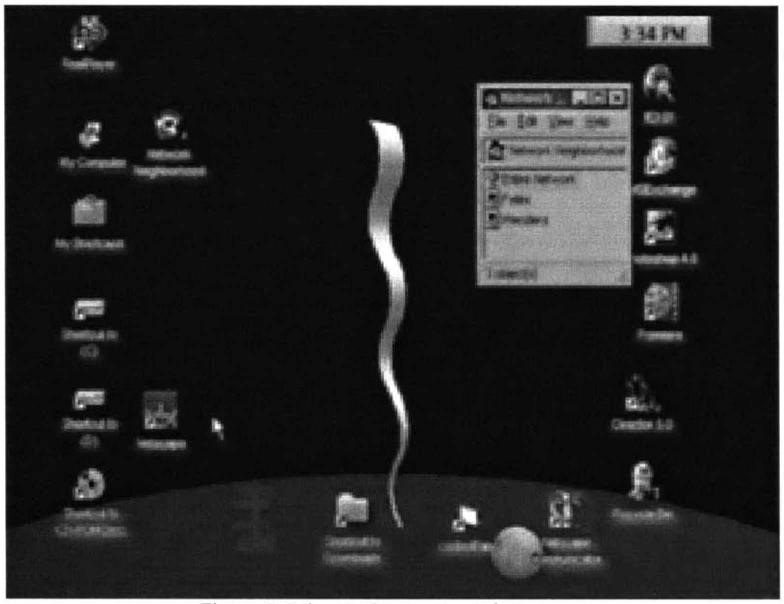

Figure 8: Mise-en-Scene, neutral state

Virtual camera positions can be changed, focussing in on the different characters for close-ups (Figure 12). An "intensity" parameter sets the camera to smoothly zoom, truck and dolly or to immediately cut to the new position, allowing control over the dramatic nature of the move. Simply being able to cut to various camera positions makes the question of drawing attention to a character much easier. A causal relation can be shown (ie: the user clicks on a link and the system cuts to a close-up), or a leisurely zoom-out to a two-shot can establish a relationship between two characters. The gaze-vector of an interface-agent no longer becomes an uncertain proposition, but can instead be clearly indicated and precisely focussed. At present, cuts and zooms can be set to correspond to jumps in processor speed: a steady increase in load can cause a zoom, and a sharp peak make the camera cut to a close-up.

A light-source is connected to system-time, adjusting its position to create environmental conditions from dusk to dawn and through nighttime. In an office without windows, these conditions can provide constant information in support of a regular clock. The brightness can also be animated, to support increased activity from the characters as needed. The opacity of the entire scene can be varied as well, desaturating and dimming along a continuous scale.

The two NURBS characters are modeled with simple movements vaguely reminiscent of plants swaying in the breeze. A gentle waving movement is the default state, with speeds and dimensions changing as needed. The rendering style can vary along a continuum from smooth and continuous to spiky and angular, implying a more jittery, anxious state of affairs (Figure 11). The b-spline surface can also take on more anthropomorphic characteristics, developing a distinct head and body to look in particular directions or at particular icons. The idea is to allow foreground and background interactions to occur with the same characters, both by making the characters more visible (ie: through scaling and color) and through alternate representations. The second character doesn't have as wide a range of motion – it communicates things like network activity, disk space and fragmentation through its general appearance. When the focus of the user is needed, faster motion and size-changes are used.

To date, the system works as a collection of visual techniques. The "intelligence" displayed is the same as with the previous systems: network traffic, processor load or other indicators of machine state can be connected to any of the environmental parameters. The various production qualities can be varied to observe effect, but aren't connected to more sophisticated software agents.

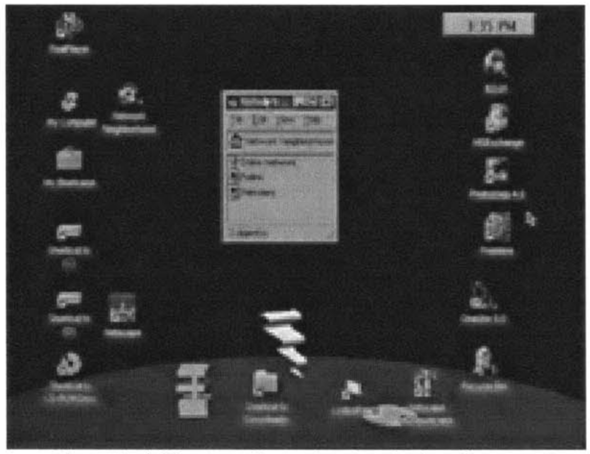

Figure 9: Affine transformations and animations

Analysis

The impact of a movie-scene is produced by many elements. If character and drama can inform computer interface, so too can set-design, lighting, and cinematography. The techniques used in this sketch aren't complex, and aren't strongly connected to specific messages or agent-intelligence. For the most part, the intention was to understand how the techniques might all work together, rather than to push any one concept very far.

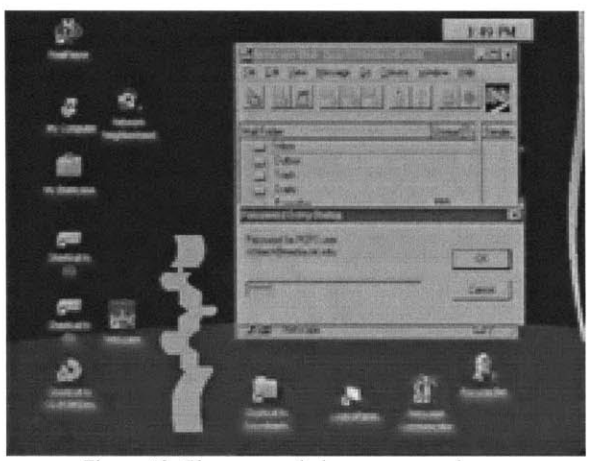

Figure 10: Close-up and character growth

With that in mind, there are several problems with the system as implemented. Abrupt scene changes (ie: cuts to and from close-ups, sudden changes in lighting conditions) are used to indicate dramatic system shifts or important messages. If the user switches rapidly between tasks, different messages become important and the character-state must also shift to compensate. The meaning of a camera-move becomes confusing when it can be linked to different tasks. It seems useful to work with a purposeful ambiguity, keeping background messages vague in order to convey many possibly-useful events rather than a few high-priority ones. More recognizable foreground techniques should be used for very specific messaging.

"Mise-en-scene" tries to capitalize on unused display space, rather than tiling window upon window, or gauge upon gauge. The desktop-background is large, pops into view when switching between several different programs, and is typically peripheral to the user's main focus of attention. The user doesn't have to think to check up on some diagnostic – it's located within an already visible niche. Having the characters be part of the scene makes foregrounding certain information difficult. Individual characters like Cypher's "Eager" were able to pop in front of windows when they had a suggestion to make or a message to give [Cypher91]; the Mise-en-Scene plants are much more closely bound to a complete environment, and loose this ability. The idea that the "teammate" characters can use and refer to the same windows as the user is lost because the link between character and setting is too strong. Another less-bound character is needed to adequately deal with this kind of messaging.

Varying the lighting conditions produces an interesting effect, with many possible extensions. Shadow mapping can be added to the characters and environment, to make the effects of the "sun" more visible. The weather, as polled from a network service, could also affect the on-screen environment: light rainfall, harsh winter light, partial cloudiness and other situations can all be implied as general impressions rather than specific data.

The system allows the emotional impact of a message to be more closely tailored to its dramatic meaning. Even software which does a good job of knowing when to alert a user to problems and when to leave them alone (ie: Norton Disk Doctor [Symantec]) can benefit from an environmentally embodied interface. The "intensity" of a warning or error can be conveyed more clearly by passive means than, say, an error number or a reference to a manual page. By thinking of the visible display as an avatar for the computer – as composed of various elements that can display state, health and "intention" -- the computer is using familiar terminology for communication. The complex system that is a modern computer can be modeled generally, abstractly, and simply. If warranted, the detailed information can be made available in a more manipulatable form. Like a graph, the system presents a visualization of information; however, in this case it uses cinematographic design elements (rather than graphical ones) to help it communicate dynamic information with greater impact.

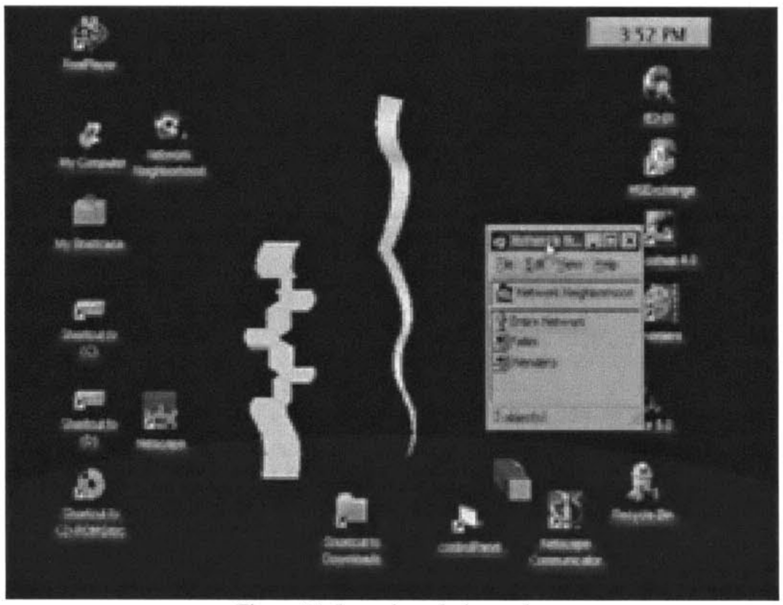

Figure 11: Jagged rendering style

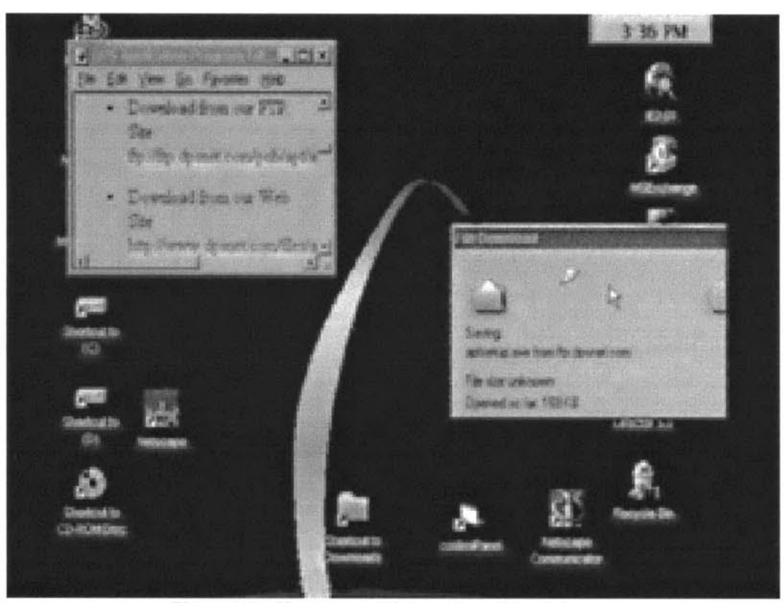

Figure 12: Close-up, anthropomorphic state

chapter five

Conclusions

Future Directions

This thesis focussed on extending the potential vocabulary of the computer interface, rather than developing a completely functional animist system. While several projects were produced which illustrate the ideas, they are not sufficiently polished to warrant regular use. More rigorous user-feedback on the techniques would no doubt generate interesting observations and refinements on the proposed ideas. This would require the development of a more functional prototype and a user-test. "GuardDog", a proposed scenario/study, is detailed below.

GuardDog

Using current technology, information about security problems in a Windows NT environment could be conveyed through animist techniques. Active-X components have access to Internet Explorer internal messages, and can intercept and re-interpret the browser's security-status reports. Net security is an area in which the computer takes a lot of action autonomously, without any way for the user to keep track of what's going on. Once a plug-in is installed, for example, there is no way to tell when a particular web-page is using it, or how.

As this information isn't always important, a system which simply requests user approval before each transaction would not be appropriate; indeed, current systems' "Yes/No" modal dialog boxes are used so frequently that the actual messages often go unread.
Desktop background color and a series of sound-samples could be used to create a guard-dog embodiment for tracking security issues (ie: submission of form-data, encryption levels, plain-text transmission of passwords, use of a secure server, automatic installation of software from "trusted" sites, downloading Java applets, use of potentially dangerous plug-ins). As the situation changes, the system can growl, bark, howl or whimper with a variety of

volumes and styles. As endlessly looping audio can quickly become annoying, the colour-changes can reflect and reinforce the audio-cues, and serve as a more continuous reminder of state.

As they surf through a variety of pre-selected or created sites, users can be asked to rate their sense of comfort and security. These can then be compared with a control-group surfing with a browser's default security warnings. If among the sites we plant several which do in fact breach security, some sense of the usefulness of the system can be studied.
In the style of Reeves and Nass [Reeves97] it may also be worthwhile to give the subjects some secondary task (perhaps related to the actual content on the surfed site) to test how distracting the security system is.

Lessons and Questions

Eventually, techniques like those in GuardDog and those discussed in this thesis can be used to make the computer-world more rich, fulfilling, and fun, as well as easier to understand. The "disembodied" character of this study can be extended into a visually represented one, allowing user-testing of camera techniques in character-based communication. Still major questions need to be answered, on topics ranging from technique to philosophy. A clearer understanding of distraction versus utility must be gained, in order to bring cinematic techniques from the pre-scripted realm to an interactive one. Cinematic language is familiar to us, but that familiarity is based on someone else making the decisions.
How an animist system can find a working balance between user-control and system-control is still unknown.

It also is interesting to research if and where these techniques can be incorporated in direct-manipulation interfaces, instead of just agent-driven ones. If animist interfaces are truly a useful means for communicating large bodies of data rapidly, then users themselves may want to make use of them in authoring or inter-personal computer-mediated conversations. Many of the principles, such as variation of background color or window-border styles, could be used to supplement existing user-interface models. Subtle variations could imply additional information when transmitting any traditional message.

The theories that came out of the iterated design/sketching process seem worthwhile to pursue further as useful principles for new styles of display and interaction. "Animist Interface" incorporates intelligence and dynamism into screen and its interface elements; it also increases in-band communication, allowing existing messages and design elements to convey additional information. Background and foreground communications styles, supported by virtual lighting techniques, allow users to maintain a better model of the state of the net, their system, and semi-autonomous software. Virtual camera moves and variations in rendering can add to the communicative abilities of animated characters, allowing the intelligence behind them to be expressed more clearly, efficiently, and enjoyably. Lightweight characters require less overhead to "parse" and understand, and may represent intelligence and functionality more accurately than overly human-like representations.

As the capabilities and functionality of computer systems increase, so too does the amount of time we spend in computer-mediated environments. It seems worthwhile to evolve those environments in ways we know to be expressive and rich, as we move towards new and unique ways of communicating.

Bibliography

Ball, Gene et al. (1996). "Lifelike Computer Characters: The Persona Project at Microsoft". *Software Agents* (Jeffrey Bradshaw, Ed.). AAAI/MIT Press. Spring.

Bates, Joseph. (1994a). "The Role of Emotion in Believable Agents". *Communications of the ACM, Special Issue on Agents.* July.

Bates, Joseph. (1994b). "Call for participation". *Working Notes: AAAI Spring Symposium Series, Believable Agents.* Stanford, California.

Bledsoe, W. (1986). "I Had a Dream: AAAI presidential address". *AI Magazine*, 7:1.

Blumberg, Bruce. (1996). "Old Tricks, New Dogs: Ethology and Interactive Creatures". PhD Thesis, Massachusetts Institute of Technology.

Borkenau, P. (1993). "How accurate are judgments of intelligence by strangers?", Presented at 101st Annual meeting of American Psychological Association, Toronto, Canada. August.

Brannigan, Edward. (1992). *Narrative Comprehension and Film.* Routledge: New York, NY.

Brennan, Susan. (1994). "Interface Agents." Unpublished paper.

Brennan, Susan. (1990). "Conversation as Direct Manipulation: An Iconoclastic View." In *The Art of Human-Computer Interface Design*, ed. Brenda Laurel. Addison-Wesley: Reading, Mass.

Buxton, William. (1990). "The 'Natural' Language of Interaction: A Perspective on Non-Verbal Dialogues." In *The Art of Human Computer Interface Design*, ed. Brenda Laurel. Addison-Wesley: Reading, MA

Clanton, Chuck and Young, Emilie. (1994)."Film Craft in User Interface Design" *SIGCHI'94 Tutorial Notes.* Summer.

Cypher, Allen. (1991). "Eager: Programming Repetitive Tasks by Example." *Proceedings of the CHI'91 Conference on Human Factors.* Association for Computing Machines: New York, NY

Denning, Peter. (1982). "Electronic Junk". *Communications of the ACM.* Volume 25 Number 3. March.

Dijkstra, E.W. (1989). "On the Cruelty of Really Teaching Computing Science", *CACM 2(12).*

Don, Abbe. (1990). "Narrative and the Interface." *In The Art of Human-Computer Interface Design*, ed. Brenda Laurel. Reading, Mass: Addison-Wesley

Elliot, Clark. (1995)."Getting to the point: Emotion as a necessary and sufficient element of story construction". AAAI Spring Symposium on Interactive Story Systems '95. Online. Available: http://www.depaul.edu/~elliott/ar.html

Elliot, Clark. (1997). "Hunting for the Holy Grail with 'emotionally intelligent' virtual actors." ACM's Intelligence97 conference. Online. Available: http://www.depaul.edu/~elliott/ar.html

Fujitsu Interactive. (1997). "Fujitsu Interactive unveils world's first CD-Rom Title incorporating Artificial Life Technology". Fujitsu Press Release. 24 February.

Gibson, William. (1996). *Idoru.* G.P. Putnam & Sons: New York, NY.

Gonzalez, Cleotilde. (1996). "Does Animation in User Interface Improve Decision making?" in *CHI'96 Conference Proceedings in Human Factors in Computing Systems.* Addison-Wesley: Reading, MA.

Grassmuck, Volker (1995): "From Animism to Animation", Byline (Dentsu Sôken). March. Online. Available: http://race-server.race.u-tokyo.ac.jp/RACE/TGM/Texts/by-line.html

Heckel, Paul. (1984). *The Elements of Friendly Software Design.* Warner Books: New York, NY.

Hoffner, C. and Cantor, I. (1991). "Perceiving and responding to mass media characters." In *Responding to the screen: Reception and reaction processes*, ed: Bryant, J. and Zillman, D. Lawrence Erlbaum Associates, Publishers: Hillsdale, N.J

Ibister, Katherine. (1995)."Perceived intelligence and the design of computer characters". Presentation to the Lifelike Computer Characters conference. Snowbird, Utah. 28 September. Available: http://www.cyborganic.com/people/kath/lccpaper.html

Keller, E.F. (1985). *Reflections on Gender and Science.* Yale University Press: New Haven, CT

King, William. et al. "The Representation of Agents: Anthropomorphism, Agency and Intelligence." *CHI'96 Human Factors in Computing Systems.* Addison-Wesleyh: Reading, MA.

Koda, Tomoko. (1996). "Agents with Faces: A Study on the Effects of Personification of Software Agents". MS Thesis. Massachusetts Institute of Technology.

Kozierok, Robyn. (1993). "A Learning Approach to Knowledge Acquisition for Intelligent Interface Agents." MS Thesis. Massachusetts Institute of Technology.

Lachman, Richard. (1997). "Experiments in Mapping Character Animation to Computational State", to be published in Workshop on Animated Interface Agents, IJCAI '97. Nagoya, Japan. August.

Lanier, Jaron. (1995) "Agents of Alienation", Interactions, Vol. II.3, July. Online. Available: http://www.voyagerco.com/consider/agents/jaron.html

Lashkari, Y. et al. (1994). S "Collaboartive Interface Agents." *Proceedings of the National Conference on Artificial Intelligence.* MIT Press: Cambridge, MA

Laurel, Brenda. (1990). "Interface Agents: Metaphors with Character" *in The Art of Human-Computer Interface Design* (Brenda Laurel, ed.). Addison-Wesley: Reading, MA

Laurel, Brenda. (1991). *Computers as Theatre.* Addison-Wesley: New York, NY

Lettice, John. (1995). "Bob – Microsoft take evasive action". The Register. 16 January. Online. Available: http://www.hubcom.com/magee/reg11.txt

Maes, Pattie. (1995a)."Artificial Life meets Entertainment: Lifelike Autonomous Agents". *Special Issue on New Horizons of Commercial and Industrial AI, Communications of the ACM*: Vol. 38, No. 11. November.

Maes, Pattie. (1995b)."Intelligent Software". *Scientific American Special Issue on Key Technologies for the 21st Century.* September.

Negroponte. (1996). "Agents: From Direct Manipulation to Delegation". *Software Agents* (Jeffrey Bradshaw, Ed.). AAAI/MIT Press. Spring.

Perlin, Ken and Goldberg, Athomas. (1996). "Improv: A System for Scripting Interactive Actors in Virtual Worlds". *SIGGRAPH'96 Course Notes #25.*

Rankin, P. J. (1993). "Research into Content Generation – Interactive Drama". Presentation to IEE Colloquium on Interactive Television. 2 October.

Reeves, Byron and Nass, Clifford. (1996). *The Media Equation.* CLSI Publications.

Selfridge, O. G. (1959). "Pandemonium: A paradigm for learning." *Proceedings of the Symposium on Mechanization of Thought Processes* (Blake, D. V. and Uttley, A. M., ed.). National Physical Laboratory. Her Majesty's Stationery Office: Teddington, United Kingdom.

Sheth, Beerud. (1994). "A Learning Approach to Personal Information Filtering" MS Thesis. Massachusetts Institute of Technology.

Shneiderman, Ben. (1997). "Advanced Graphic User Interfaces: Elastic and Tightly Coupled Windows." Online. System Concepts Limited. Available: http://www.system-concepts.com/masterclass/benpaper.html

Shneiderman, Ben. (1996)."Direct manipulation vs. Agents: Paths to predictable, controllable, and comprehensible interfaces". *Software Agents* (Jeffrey Bradshaw, Ed.). AAAI/MIT Press. Spring.

Shneiderman, Ben (1995). "Looking for the bright side of user interface agents." *ACM Interactions*, vol. 2, # 1 (January).

Stephenson, Neal. (1993). *SnowCrash.* Bantam Books: New York, NY.

Sterling, Bruce. (1995). "Lifelike Characters". Speech to the Lifelike Computer Characters conference. Snowbird, Utah. 29 September. Online. Available: http://www.research.microsoft.com/research/ui/lcc/sterling.htm

Takeuchi, Akikazu. et al. "Situated Facial Displays: Towards Social Interaction" *CHI'94 Conference Proceedings; Human Factors in Computing Systems.* Addison-Wesley: Reading, MA.

Travers, Michael. (1996). "Programming with Agents: New metaphors for thinking about computation". PhD Thesis, Massachusetts Institute of Technology.

Thomas, Frank and Johnston, Ollie. (1981). *Disney Animation: The Illusion of Life.* Abbeville Press: New York, NY.

Walker, Janet. et al. "Using a Human Face in an Interface," *CHI'94 Conference Proceedings; Human Factors in Computing Systems.* Addison Wesley: Reading, MA

Williams, Margery and Green, Michael (Illus.) (1984). *The Classic Tale of the Velveteen Rabbit: Or, How Toys Become Real*. Running Press: Philadelphia, PA

Wernecke, Josie. (1994). "The Inventor Mentor." Addison-Wesley: Reading, MA.

URLs and Videos

"Animism". (1997). Britannica Online. Online. Available: http://www.eb.com 27 July.

Chapter One. (1988). Apple Computers. Videotape.

Creatures. (1997). Cyberlife Inc. Online. Available: http://www.cyberlife.co.uk/welcome_frameset.htm

"Firefly". Firefly Net. Online. Available: http://www.firefly.com

Knowledge Navigator. (1987). Apple Computers. Videotape.

Microsoft Bob Microsoft, Inc. Online. Available:

Microsoft Agent Microsoft, Inc. Online. Available: http://www.microsoft.com/intdev/agent

Microsoft Office 97. Microsoft, Inc. Online. Available: http://www.microsoft.com/office/

Norton Disk Doctor. Symantec, Inc. Online. Available: http://www.symantec.com/nu/fs_nunt.html

PFMagic Homepage. PFMagic, Inc. Online. Available: http://www.pfmagic.com

"What About Animism?". (1995). ICT, Inc. Online. Available: http://www.ict.co.jp/japa-bus/animism.html

Wohl (1995) "Social Graces, Social Interfaces". Wohl Associates. Online. May. Available: http://www.wohl.com/wa001.htm

Lecutures and Conversations

Maes, Pattie. (1996) In "MAS334/MAS737: Software Agents Seminar". 25 November.

Reeves, Byron and Nass, Clifford. (1997). Media Lab Colloquium, 29 May.

Wexelblatt, Alan. (1997) Personal conversation. March