

# THE HUMAN AVATAR PROJECT

Thesis Document  
May 10, 2007

Ian Spangler  
Candidate for MFA, Design & Technology  
Parsons The New School for Design  
[iaspangler@yahoo.com](mailto:iaspangler@yahoo.com)  
<http://ian-spangler.com>

Instructors: Ted Byfield, Stephanie Owens, Louisa Campbell  
Advisor: Igor Pusenjak  
Thesis URL: <http://a.parsons.edu/~ispangler/thesis>

# Chapters

<b>1: Concept and Design Questions .....</b>	<b>3</b>
<b>2: Domains and Precedents .....</b>	<b>13</b>
<b>3: Methodology .....</b>	<b>25</b>
<b>4: Evaluation .....</b>	<b>47</b>

# 1: Concept and Design Questions

## 1.1 Motivation

Today's computer games and online virtual experiences featuring human avatars provide users with a high level of immersion and control through their 3D virtual simulations, but for all of the options and "add-ons" they provide, they seem to be unable to dig beneath the surface. As they cater to increasing the user's arsenal of choices, astute players will realize that something is missing. When the discriminating user zooms in close and stares into the face of the fabricated polygonal manikin that is supposed to represent him, an overwhelming void and lack of meaning is felt. At best, the human avatars of current simulations are an impressive extension of a person's physical presence in the world. However, they feel alarmingly disconnected from the user's inner self.

The majority of avatar simulations revolve extensively around physical appearance and control of a body. Digital worlds such as *Second Life* (Linden Labs) and *The Sims* franchise (Maxis/EA) allow users to create a detailed three-dimensional likeness of the user, one that mirrors the user's appearance or the appearance that the user would like to have, down to a tattoo on the arm. A large part of creating a *Meez* avatar, a "3D I.D." of a web user, involves shopping for clothes items in an extensive wardrobe gallery. The user can precisely manipulate the created body like a puppet, using onscreen buttons and a keyboard, using it to navigate a house, a terrain, or a club in cyberspace and present itself to the puppets of other users. This puppet, soulless and robotic, obeying the user's every command, has come to be known as an avatar.

Avatars, at least those that are in some way humanoid, seem to have their roots in the concept of a character. Characters are the agents of stories, but the person reading a book or watching a film does not interact with or control a character. A character affects and is affected by the fictional world that it occupies in a way that is predetermined by the story's author. Watching or reading about a character lacks much of the stimulation that real-time interaction provides, but movies and books do have certain

experiential qualities which games lack. In particular, the nuances of a pre-imagined script and the emotive powers of actors allow movie watchers to feel as if they are immersed in the mind and the feelings of a character. Most games involving human characters or avatars only attempt to give players an adrenaline rush, a very rudimentary form of entertainment, with a bit of puzzle solving thrown in.

This author's interest in interactive entertainment has evolved out of a background in film and video theory, production, and visual art. As games become more real and life-like, employing humans as characters or avatars, they inevitably face a more pressing comparison with the older and more mature medium of film. Attempting to mimic film in its technical verisimilitude is a worthy cause, but it is only half of the battle. In film lies a richness and depth that has evolved over many decades. Games employing virtual actors have far to go in obtaining these nuances, but introducing a new angle to interactive experiences, one that is not solely concerned with outward appearances or the all-mighty control over a body, may help steer them in a more meaningful direction.

## 1.2 Goals

Based on these observations, the main intention is to investigate and devise alternative modes of representation of a human avatar to reveal more about the human mind and emotions in the experience of engaging with one of these simulations. Providing an alternative means for an avatar to be represented in relation to its user may be the key to creating an experience that is instructive and rich in metaphor in the way that films can be.

A fundamental part of making the user feel more connected to his or her avatar is not to give the user more control over its body, but instead to introduce a greater level of complexity to their psychic relationship. On a physical/ perceptual level, an avatar occupies a unique territory of identity somewhere between self (unseen first person navigator) and other (character). The goal of this thesis is to extend this hybrid form of identification to the mental-emotional level, fostering a more complete fusion of consciousness.

### 1.3 Driving Questions

In developing an avatar simulation that supports a new kind of user-avatar relationship, one in which users are encouraged to identify with something larger than the body and its appearance, a primary question is how the mind and emotions may perform a key role in the simulation. How can the mind of the user/avatar be represented? This is a nebulous area, one from which many designers may shy away because it inevitably raises the unfathomable question, whose mind? Is it the user's mind or the avatar's? Can the avatar appear to have its own mental processes even though its mind is not completely its own?

What about multiple users? Can multiple users share the same avatar? Since people are very used to localizing identity in a physical body, this is an interesting proposition. How might a number of users be able to simultaneously affect a single avatar without stepping one another's shoes?

To answer the above questions, a designer may need to find a way to externalize the thinking process — to visualize the thoughts of the user/ avatar on screen. In addition, there is inevitably the question of actual content to engage the user or player. Creating a meaningful interactive meditation on the mind and the emotions requires that the designer concretize it in the form of a story or scenario.

On a technical and aesthetic level, the increasing malleability of video as a web-based medium makes it an intriguing alternative to 3D animation as a means of visually representing an avatar. However, the goal is not to produce a linear film or even a video with a branching narrative structure. The strengths of avatar simulations lie in their artificially intelligent and emergent qualities. Hence, it is desirable to maintain a relatively high level of changeability and dynamic interaction in the proposed simulation.

### 1.4 Initial Concepts

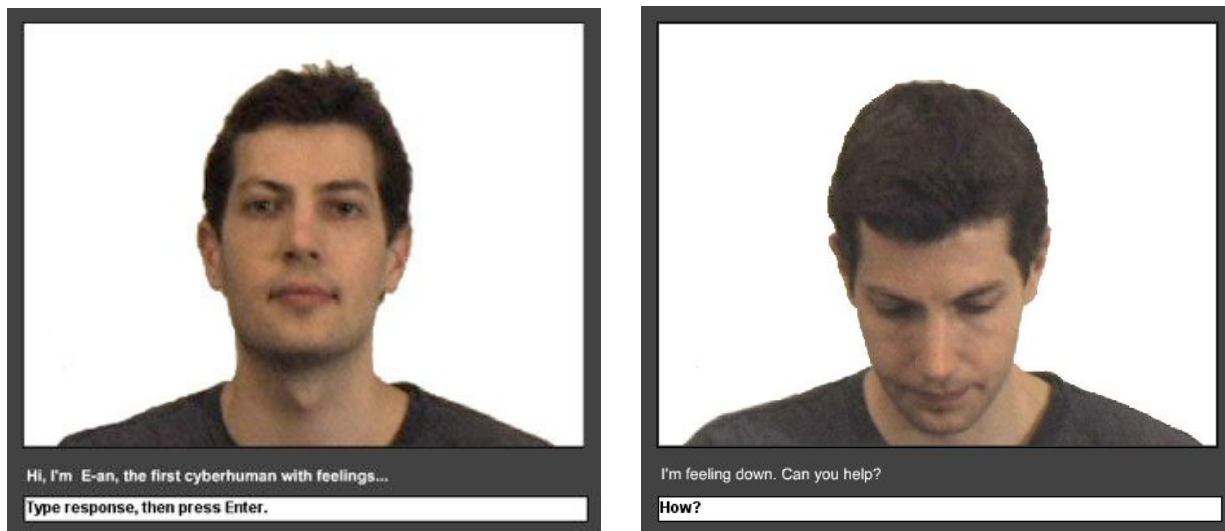
At the outset, the intention was to design an experience around a fundamental dynamic struggle of the human mind. The battle between positive and negative thinking (resulting in positive and negative life habits) is one such universal dynamic.

#### 1.4.1 *Editing Joe Schmoe's Mind in a Wiki*

An initial idea was to have multiple users edit the thoughts of an average Joe in a Wiki. Changing the thought patterns would cause changes in a real-time video stream of Joe as he goes about his daily routine. He would either develop positive or negative habits depending on the ascribed connotations of particular keywords in the text.

#### 1.4.2 *The Emoting Multi-user Chatterbot*

A second concept was an emoting multi-user chatterbot. A chatterbot is a type of web-based avatar that is programmed to respond “intelligently” to user input, such as text. The idea was to create a unique type of “chatterbot” that interacts with multiple users over the Internet and retains memories of its conversations in the form of keyword-triggered emotions. The chatterbot’s behavior and dialogue would emerge in a variety of ways as it evolved emotionally over time. (Fig. 3) If users entered too many negative or disparaging words when conversing with the avatar, he would grow more angry or depressed and perhaps seek out therapy from future users.



**Fig. 1** The emoting multi-user chatterbot

#### 1.4.3 *Life Sentence.* (version 1)

Another, more developed, concept was an exploration of the mind’s affects on the fate of the human body. This turned out to be a rather pointed criticism of the daily American lifestyle, particularly

the pervasive problems of eating fast food, overworking and alcoholism, collectively contributing to the ill health of the average American. The idea was that these negative habits were the result of negative thought patterns, and each negative habit was visually associated with a concrete symbolic piece of architecture in a suburban cloudscape such as a daunting office building or a fast food restaurant. (Fig. 1)

The avatar's body would evolve physically in the "World View", getting fatter and sicker from falling into negative habits, and getting leaner and healthier from following positive habits. However, a satisfactory mental-emotional component was missing. Original mockups of the "Mind View" for this proof of concept did not show the avatar nor did they portray his emotions. (Fig. 2)

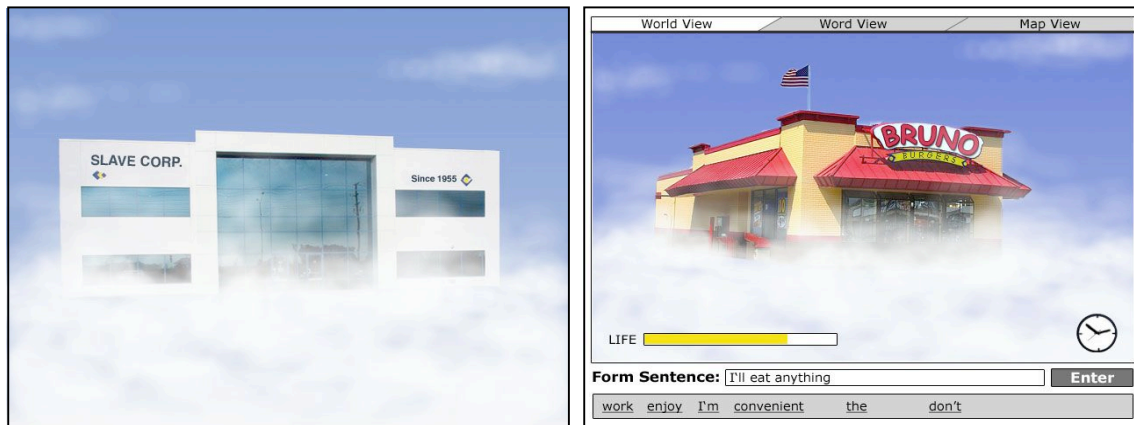


Fig. 2 Mockups of the "World View" in the first version of *Life Sentence*.

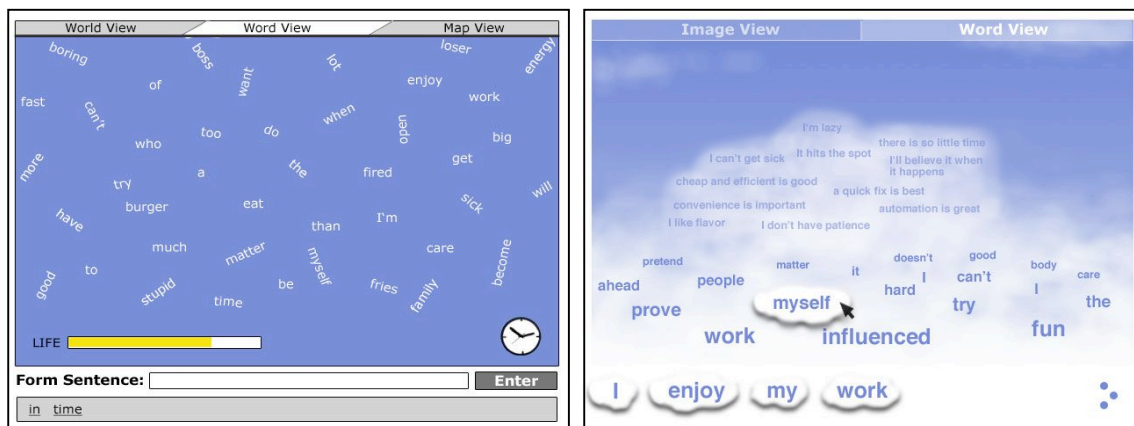


Fig. 3 Mockups of the thought formation process in two different iterations of *Life Sentence*.

#### 1.4.4 *Life Sentence.* (version 2)

In a second version of the *Life Sentence.* concept, the qualitative nature of the mind and the emotions, as opposed to physical health, form the crux of the product's message, and each life structure represents a psychological "chamber of experience." *Life Sentence.* was subsequently tagged as a web-based virtual experience/ game in which multiple users channel their collective consciousness into the mind of a single transpersonal human being. Exploring the differences between the effects of positive and negative thought patterns on a person's inner and outer life, the application features a shared avatar that evolves emotionally based on the types of thoughts it receives and the experiences that are generated from these thoughts.

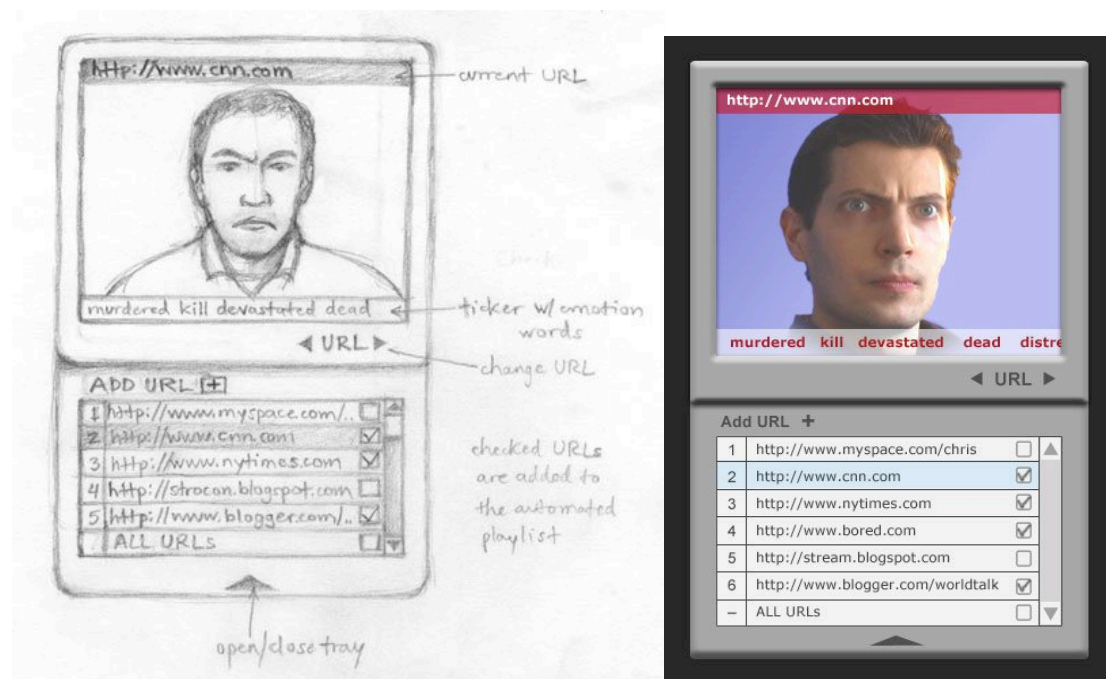
Users would drag words from a cloudstream into the avatar's mind bubble to form thoughts, the emotional connotations of which would update his facial expressions. Some thoughts would bring a certain life structure (home or work for example) into existence, and the avatar's emotional state would contribute to his having a particular kind of emotional experience, negative or positive, within the life structure. This in turn would change the emotional composition of the cloudstream and cause certain thought patterns or mental traps to emerge that the user must try to avoid. The potential end result would be that the avatar would hurl himself off the edge of a bluff of clouds due to a severe negative emotional state that is allowed to persist for too long. (For a full description of this concept, see *Appendix*)



**Fig. 4** Initial pre-visualizations of the 'Mind View' and 'Life View' in the second version of *Life Sentence.*

## 1.5 Final Concept

The final concept, *Universal Emote*, represents a nearly complete departure from a representation of an external world. Instead it hones in on the essential relationship between the avatar's mind and emotions as the product of mass consciousness on the World Wide Web. *Chap 3: Methodology* provides a detailed discussion of how prototyping thought-formation interaction led to this final concept.



**Fig. 5** A Sketch and Full Graphical Mockup of the *Universal Emote Control* interface

*Universal Emote* is the brand name for a unique suite of web-based applications that visualize the emotional effects of collective thought patterns on a relatable human face. One of the products in the suite, the 'Universal Emote Control', is a downloadable widget that tracks the emotional content of web pages in the form of facial expressions. As users create and manage their own list of URL "channels", the program uses psycholinguistic data to analyze the overall affect of the text in each URL and continually provide updates. The Universal Emote Control has been prototyped to high degree of fidelity and represents a proof of concept. (See *Chap. 3: Methodology*)

In addition to the widget, the promotional web site for *Universal Emote* also features an in-browser version of the Emote Control, one in which multiple users contribute to a growing list of URL channels. *Universal Emote* games include a “head-to-head” multi-player competition in which each player tries to emote his/ her own version of the avatar to the height of a particular emotion before all of the other players (by typing URLs), and a *therapy game* in which the user plays therapist, trying to guess keywords that may be contributing to the avatar’s negative emotional state based on a list of URLs that are infusing his consciousness. (See *Chap 4: Evaluation* for an in-depth explanation of future directions)

*Universal Emote* applications and games are designed to entertain and enlighten the casual web user by channeling mass psychological processes into a personal form. The video-based avatar they feature is built using a special technique of blending three separate facial components seamlessly together. It is designed on a model of four primary emotions that can mix dynamically in a variety of ways depending on the emotional composition of the text it takes in.

The decision to eschew a digitally spawned avatar effectively eradicates associations with conventional games such as *The Sims* (Maxis, 2000) and *The Sims 2* (Maxis, 2004). Affecting a “real” human being, as opposed to a cartoon, on an emotional and mental level will likely have a more visceral impact on users. There is potentially a rather haunting, but pleasurable feeling to be derived from putting thoughts and emotions into the husk of an individual that is known to be existent in the real world but whom has given up his identity in a virtual world.

## 1.6 Philosophy

*Universal Emote* stream-of-consciousness applications constitute a psychological framework that is universally applicable to all human beings. They explore the interrelationship between thoughts, emotions, and language from a perspective informed by a mixture of Eastern, Western, and New Age thought. If there is any primary objective to these types of experiences, it is to promote a certain degree of mindfulness or attentiveness to one’s thoughts, a valuable exercise espoused by both Buddhists and some cognitive psychologists. The user is posited simultaneously as a participant in the stream of consciousness

and as a “watcher, ‘the awareness that watches and responds to the meanings that appear in the stream of consciousness.’” (Ekman, 73-74)

### 1.6.1 Zen

One of the driving principles behind the project is an outgrowth of Zen Buddhism. By positing the mind as the source of emotional vicissitudes, it rejects the notion that outward experiences or stimuli alone account for a person’s state of wellbeing. The simulations present a model of the reflective, rather than the functional, mind. The reflective level of the mind consists primarily of thoughts that either trigger emotions or have no actual value other than to create clutter.

### 1.6.2 Carl Jung’s Collective Unconscious

Swiss psychiatrist and pioneer of analytical psychology Carl Jung devised the term “collective unconscious” to refer to archetypal patterns or symbols in the psyche that are ingrained in man’s primordial nature and hence common to all human beings. Examples include the hero, the anima/animus (masculine/feminine), the mother, and the shadow (repressed secrets). The relationships between these archetypes become enacted in our life dramas. While the *Universal Emote* applications do not specifically draw from any of Jung’s archetypes or deal with parts of the mind that are completely inaccessible, they do reaffirm the notion that the way one thinks is not individually determined. It is a collective psychological predisposition. (Hall, 38-55)

## 1.7 Target Audience

*Universal Emote* applications are aimed at a broad audience of casual web users but it may be enjoyed most by one of the following types of individuals:

- The *Sims* gamer looking for a new psychological dimension to the human avatar experience
- The casual gamer looking for a mix of metaphor and entertainment beyond the mechanics of a puzzle or psychomotor action simulation

In particular, the following types of users are likely to take an interest in the Universal Emote Control widget:

- The web user seeking out the latest smart Internet gimmick, tech-fad, or gadget
- The web user who tends to cognitively dissociate from his/ her emotions and harbors a subconscious need to be “emotionally informed”

## 2: Domains and Precedents

### 2.1 Background on Related Work

This project calls together a variety of technological design domains, including psycholinguistic analysis, low-level artificial intelligence, Internet data mining, and interactive video. The experience fuses these domains without resting too heavily on the foundation of any single discipline.

### 2.2 Connotative Meanings in Language

As words are channeled into the avatar's head in a *Universal Emote* application, user witness a particular causal relationship between the text and the expressions exhibited on the avatar's face. Collectively, the words trigger new emotions or enhance existing ones. Hence, a significant aspect of the project involves associating a large number of words and phrases in the English language with certain emotional, or connotative, meanings. This task has typically fallen under the umbrella of psycholinguistic content analysis. A variety of programs, some web-based, have been developed to procedurally scan user-entered text, evaluating the writing statistically on a number of semantic factors including the evocation of positive and negative feelings. Most often, the programs make these assessments by looking up keywords and phrases in a database or dictionary and retrieving associated values that have been validated through research studies. In ascribing universal connotative meaning to bits of language, emotional content remains the most elusive of all semantic factors because of technical complications like part of speech and context.

#### *Charles Osgood's Semantic Differentials*

A pioneer in the study of subjective meaning in words and in "objects" in general was the psychologist Charles Osgood. In the 1950s, Osgood devised a scale for plotting adjectives along a continuum called the semantic differential, or semantic space. (Underwood, 2005) Through a number of

surveys, he evaluated the words along 7-point scales according to three primary attitudinal factors: Evaluation (how people rate the word according to certain qualitative polarities such as ‘good’ and ‘bad’, or ‘cold’ and ‘hot’), Potency (how people rate the word along the spectrum of ‘strong’ vs. ‘weak’), and Activity (how they rate the word along the spectrum of ‘active’ vs. ‘passive’) (Jay, 415). Osgood’s work was the first to give statistical weight to connotative word meanings, but it was not able to ascertain the emotional content of positively or negatively valenced words.

#### *A.N.E.W.*

Affective Norms for English Words is the product of an ongoing study at the Center for the Study of Emotion and Attention (CSEA) at the University of Florida. Their published data evaluates a large number of English words (nouns, adjectives, and some verbs) along semantic lines similar to those of Osgood. The categories include positive or negative valence (with higher positive values denoting increasing Pleasure), Dominance, and Arousal. (Bradley, 1999) These kinds of ratings have become a gold standard in psychological research, but are weakly connected to expressible emotions such as anger, sadness, fear, and joy. However, through additional analysis of the values, it is possible to decipher general emotional meanings. (See *Chap 3: Methodology*)

#### *The EMMI project*

A pioneering linguistics researcher by the name of Wayne Chase overcame the technical obstacles in determining the emotional content of words and phrases. He spent several decades developing what he called the Emotional Meaning and Impact Analyst, a piece of software connected to a connotative/ emotional dictionary that was designed to analyze writing and deliver a detailed evaluation of its overall emotional impact on a societal level. (“History of Connotative Reference Technology”, 2007) Currently, Chase heads a Vancouver group that is dedicated to developing a whole series of connotative reference tools for creative writers, speechmakers, and other types of users, including connotative equivalents of a dictionary, thesaurus, name giver, connotation checker, and others.

(“Connotative Reference Corporation”, 2007) These references remarkably catalog emotional connotations for just about every word and phrase in the English language, including slang expression and basic verbs that seem quite neutral. (“Connotative Reference Corporation”, 2007) Unfortunately, none of the tools will be commercially available for at least a couple of years.

#### *L.I.W.C.*

The Linguistics Information and Word Count, a text analysis software program designed by James W. Pennebaker, Roger J. Booth, and Martha E. Francis, evaluates keywords along a number of dimensions, one being psychological affect. (“LIWC”, 2007) Factors include positive feelings, optimism and energy, anxiety/ fear, anger and sadness/ depression. The downside to this program is that by default, it only designates 615 words as being positively or negatively inflected and hence fails to detect affect in some cases. However, a major selling point is that it enables developers and analysts to supplement and shape the data to their liking by creating a customized dictionary suitable to their own needs.

#### *D.A.L.*

Whissell’s Dictionary of Affect in Language (DAL) rates words along many of the same core dimensions as ANEW, but extrapolates certain emotions from the raw values assigned for these dimensions. Based on activation and pleasantness values, DAL designates certain words as being “sad” or “nasty.” With over 8700 words catalogued and rated, DAL reports a ninety percent return rate for text entries written in English. (“Whissell’s Dictionary of Affect in Language”) DAL could be a useful supplement to LIWC for these reasons.

### **2.3 Emergence and Artificial Intelligence**

*Universal Emote* applications are predicated on the notion of an avatar’s facial expressions being the automated creation of a multitude of words, each word being its own “agent” with its own data to contribute to the avatar’s face. In his book Emergence, Stephen Johnson distills some basic principles

about “intelligent”, self-organizing systems out of scientific observations on the behavior of ant colonies and historical data on the development of cities. According to Johnson, “the movement from low-level rules to higher-level sophistication is what we call emergence.” (Johnson, 18) With proper feedback and control, some very discernable and often unpredictable macro-patterns can grow out of uncoordinated local agents following their own simple set of rules. (Johnson, 40) Emergence is the key to simulating intelligence in a computer program.

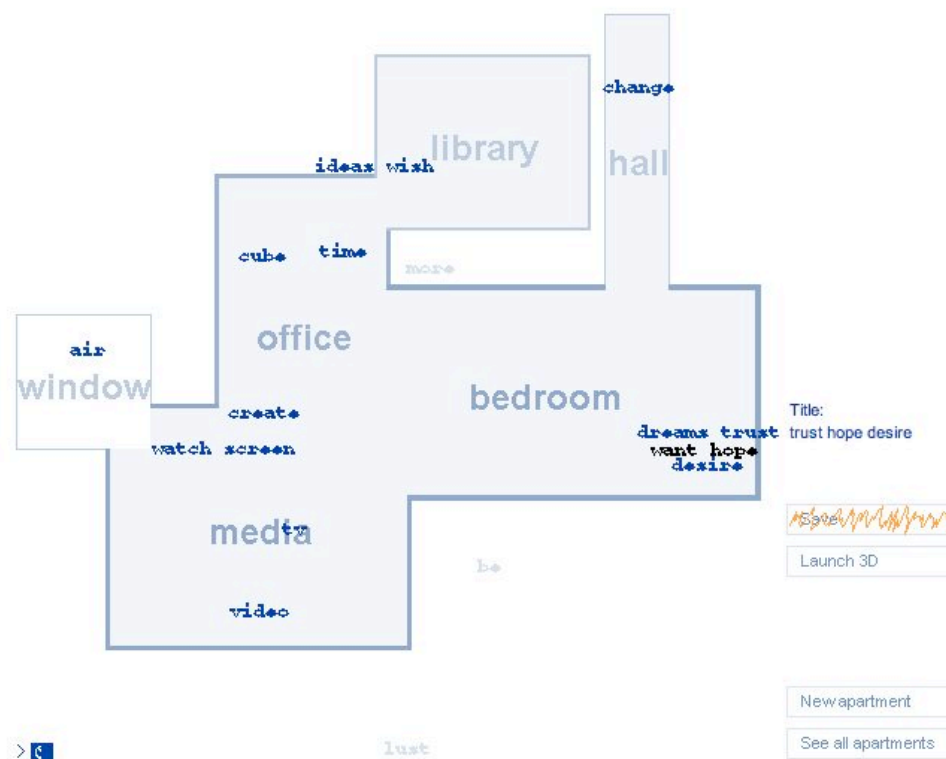
### 2.31 *The Apartment*

“Apartment”, designed by Marek Walczak and Martin Wattenberg, is a unique web-based application that, like *Universal Emote* applications, analyzes user input in the form of text and creates a larger significant whole out of the individual semantic associations of different words. (Walczak, 2002) The playful experience invites users to type a line of text at a command prompt and then watch as individual keywords float into the floor plan of an apartment, occupying and expanding a particular room with which the word is universally associated. For instance, if the user types “I’m hungry”, both of these words move into the kitchen and increase its size in relation to all of the other rooms. Words such as “sleep”, “dream”, or “sex” make the bedroom larger. Over time, the layout of the apartment becomes a meaningful picture of many users’ subjective mental states. Users can view, edit, and contribute to several growing virtual cities of these dynamically generated floor plans on the web site.

“Apartment” seems to work best when its logic feels satisfyingly “right”, when words go into the places that we might expect them to, but what really holds the user’s attention is seeing how these minor expectancies add up to something greater (a unique space or territory carved by a multitude of words) that we cannot mentally fathom at the outset. Even more interesting is how the authors were able to program the system to recognize broad macro-patterns from the all of input data it receives and present it back to the users. The diagrams of virtual apartments are visualized in a way that users can see the relationships among various categories of collective mental concerns such as Work, Glamour, Body, Truth, Food,

Intimacy, and Change along an additional dimension of time (days of the week on which users are creating these apartments).

Disappointment results only when an observed interaction between words does not seem to contribute to the meaning of the whole. For instance, as the words float through the rooms of the apartment, they continually attach and detach themselves from other words according to some inconsistent grammar rules, forming two or three word phrases. Though at times, poetic patterns of movement may emerge from this behavior, there is not any perceivable effect on the apartment itself.



**Fig. 1:** Apartment (2002) Marek Walczak and Martin Wattenberg

### 2.3.2 *Facade*

“Façade” (2005), an award-winning independent cyberdrama, operates on an even higher level of artificial intelligence. (“InteractiveStory.net”, 2005) Over the course of five toilsome years, the creators Michael Mateas and Andrew Stern developed interactive cartoon characters that function as adaptive

agents, independent of a prescribed storyline. The narrative changes moment by moment in response to the user's choices. The user talks to the characters as an unseen first person conversant by freely typing in lines of dialogue, and the program analyzes the text in fragments and determines how the characters should respond accordingly.

“Façade” is essentially a database-driven narrative that progresses dynamically, segment by segment, or one line of dialogue at a time, based on a history of user decisions. This piece sets a certain standard for avatar simulations that use text-based input.



**Fig. 2** Façade (2005) by Michael Mateas and Andrew Stern

## 2.4 Human Avatar Games and Simulations

Human avatars appear all over the world of new media. They occupy spaces as diverse as marketing web sites, console computer games, video teleconferencing, and virtual reality simulations, and they are designed to take on many different types of roles and forms in relation to their users.

### 2.4.1 *The Sims, The Sims 2, and The Sims Online*

The Sims franchise of games, including *The Sims 2* and *The Sims Online* (Maxis/ Electronic Arts, 2002, 2003), is a suite of large-scale, detailed virtual experiences that offer players the opportunity to create and micromanage the daily life of a family, or neighborhood, of human avatars. The one-to-many creator-avatar relationship, in which the user can switch from one on-screen persona to another was somewhat of a novelty when it first came out.

The Sims, as the avatars are called, function like dolls or lab rats that have had their emotions and identities cut out and placed on a table for the user to tweak. This is certainly an interesting treatment of reality — a sort of displacement of the body from psychological factors. *Universal Emote* exaggerates this effect by introducing external web feeds and multiple user input to further diffuse and delocalize identity. The idea of a multi-user stream-of-consciousness avatar may represent the first many-to-one user-avatar relationship.

The Sims games present the daily life of a variety of colorful characters from a largely extroverted and objectified viewpoint. Players are given a multitude of choices about their avatars' appearances, objects that they want to acquire, other Sims that they want to interact with, and decorative details they want to add to their houses. Conversely, *Universal Emote* games are largely introspective in nature. Whereas the *Sims Online* functions primarily as one big chat room, participants in *UE* multi-user simulations communicate indirectly, their decisions about what thought-producing URLs or words to place into the avatar's mind being implicitly influenced by the choices of others.

Perhaps most significant is how *Universal Emote* differs from Sims thematically. One of its primary messages is that activities and experience alone, or the inability to fulfill the avatar's "aspirations", are not the main perpetrators of negative emotions. Rather, thoughts, the tidbits of language that continually fester the mind, are the primary culprit. This is a quintessential distinction between Western and Eastern philosophy, and one that is deeply imbedded in the design of *UE* products.

#### 2.4.2 *Second Life*

*Second Life* (Linden Research, 2007) is a massive multi-user virtual world in which users adopt a new persona, albeit in the form of a body and its numerous material possessions. It is a vicarious experience in which people can live out dreams and pleasures that they would not ordinarily be able to live out, or would be reluctant to live out, in the real world. The world has its own currency, real estate to purchase, social clubs, and other attractions. Unlike *The Sims* games, there is not any explicit psychological or emotional dimension. All activity is centered outside of the avatar's body.

### 2.4.3 Oddcast

A commercial online company called Oddcast showcases speaking 3D and 2D avatars that are available for purchase and deployment on web sites. ("Oddcast Home", 2007) A company or group of individuals can easily lend their online presence a sense of personal embodiment with one of these avatars. In its head-centric format, the Universal Emote Control may bear the most resemblance to an Oddcast avatar display. However, Oddcast avatars are very personality-centric and focused on individuation. Interaction with one of these avatars predictably keeps the user as a separate entity from the avatar as language is outwardly directed.



Fig. 7: Oddcast avatars

## 2.5 Interactive Video

In preparation for developing a unique video-based avatar experience, it is important to consider the evolution of pre-recorded interactive video as an artform. With increasing support for rich multimedia content on the World Wide Web, interactive video has enjoyed a noticeable resurgence since it quickly fell out of favor in the mid to late 1990s. In its heyday as a laser disc technology in the 1980s, prior to the proliferation of more flexible and cost-effective 3D technologies, computer-processed video was capable of providing hyper-realistic images for users to interact with. (Horowitz, 2005) However, particularly in the case of games, interaction was often confined to “branching”, in which one video clip would transition to another according to a series of multiple-choice-based decisions.

### 2.5.1 *Full-motion Video Games (“Interactive Movies”)*

Beginning with *Dragon’s Lair* in 1983, this genre of game-making involved the use of pre-rendered live action or cartoon video sequences to create interactive stories with limited player involvement. (Horowitz, 2005)

### 2.5.2 *Video Objects and Video Avatars*

Virtual reality and tele-immersive applications treat video as a more malleable entity, segmenting “video objects” from their original backgrounds and using them in three-dimensional or mixed reality environments. Video objects derived from moving images of humans that are intended to either represent or interact with the user are called “video avatars”. They are less commonly employed than 3D avatars, but can be found across several different domains, including animated web sites, digital games, telepresence, and virtual reality. *Universal Emote* applications utilize a video avatar as opposed to a 3D model for several reasons that are discussed in *Chap3: Methodology*.

On the Internet, many web sites created with Macromedia Flash are employing video in an object-oriented fashion, blending it seamlessly with photographs, animations, and other kinds of graphics as opposed to presenting it in a standard box with a controller. A futuristic site produced for Vodafone in the UK is a potent example. (“Vodafone”, 2007) The introduction features a video avatar that confirms

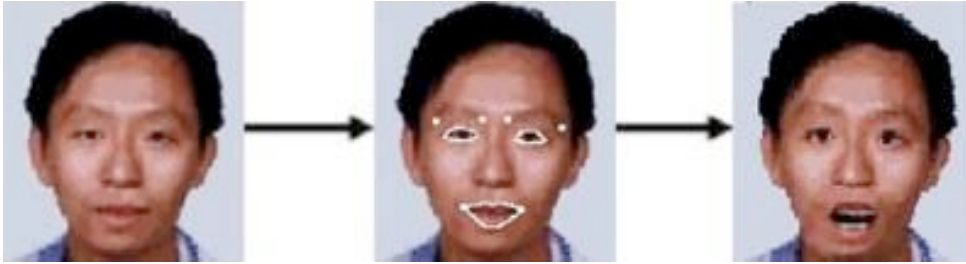
the user's selection by literally pressing the appropriate graphical button on screen. An inventive promotional site for the *Wedding Crashers*, produced in 2005, takes the idea of dynamic video even further, giving users the power to customize the video's appearance. ("Wedding Crashers Official Site", 2005) It allows users to upload a headshot and watch the movie trailer with their own face in place of the one of the lead characters.



**Fig. 4:** A) Vodafone future site avatar, and B) *Wedding Crashers* comical video generator application

## 2.6 “Expressive Real-time Communications”

A team of researchers at Bournemouth University in UK, headed by Anthony C. Boucouvalas, a professor of Communication Networks and Applications at the University of Peloponnese in Greece has developed an online text-based chat system that uses an “emoticon detection engine.” (“Research Projects”, 2007) Instead of using the typical smiley faces, users can upload a neutral photo of their own face, make some markings with their mouse, and automatically see the image converted into a series of emoticons. Through sophisticated image morphing techniques, the eyes, eyebrows and mouth are adjusted to convey one of six primary emotions (happy, sad, angry, scared, disgusted, and surprised). Each emotion has three intensity levels. (“Forget ‘emoticons’”, 2007)



**Fig. 5:** A diagram showing how a user’s face gets morphed into an expression of surprise

The researchers are working on integrating the facial morphing software with a dictionary of about 18,000 words and phrases to detect the emotional content in a message. (“Forget ‘emoticons’”, 2007) Clearly an impressive technological feat that bears some of the same components as a *Universal Emote* application, the system nonetheless emphasizes functionality over artistry and entertainment value and does not fall into the same interpretive realm as this thesis project. The interface and type of communication that the technology supports are markedly different.

## 2.6 Facial Sculpture

In 2002, an avant-garde artist named Tim Hawkinson converted a photograph of his own face into a surreal robotic relief sculpture that he called *Emoter*. (“ACE Gallery: Tim Hawkinson”, 2007) A hydraulic machine hooked up to the artist’s visage was programmed to trigger sensors that animate numerous individual facial fragments such as upper and lower eyelids, lips, eyebrows, and nose. The result is purposely more surreal than realistic, but combines principles of artificial intelligence with a photographic aesthetic quality. The similarly self-reflexive *Universal Emote* avatar aims for a higher degree of naturalism and authenticity in spite of its surreal sensibility. Nonetheless, the idea of breaking the face up into separately controlled components is one that significantly informs the process undertaken in representing the avatar’s face. (See *Chap 3: Methodology*)



**Fig. 6:** Emoter (2002) by Tim Hawkinson

## 3: Methodology

### 3.1 Overview of the Design Process

As described in Chap. I, *Universal Emote* signifies a type of web-based application that visualizes collective thought patterns in the form of facial expressions on an anonymous human face. Designing and developing this project involved numerous iterations. Variables that changed considerably in the process include the role of the user, the nature of user interaction, the form of the interface, and the conceptual framework. In the prototype's refinement from a self-contained exploratory virtual experience to a customizable widget built around validated psychological data and external web-based content, qualities of rationality, approachability, and real world contextualization superseded philosophy, exploration, and poetry.

In developing prototypes of the application, this author employed a variety of technological design tools, including Macromedia Flash/ Actionscript for the interface and user interaction, Macromedia Coldfusion and Microsoft Access for the database, a video camera and blue screen for recording the avatar's facial expressions, and Adobe After Effects for compositing the facial components and preparing them for manipulation in Flash.

### 3.2 Representing Thoughts as Words

At the outset, the goal was to create a virtual experience that externalized the notion of a thought process, highlighting and bringing awareness to certain universal patterns of the human mind by making the user a conscious "thinker". The relationship between the user's thought patterns and emotions was to become a focal point of the interaction. In particular, the aim was to provide an accessible model of the *reflective* component of the human mind, the part of the mind that evaluates, judges, and conjures up or strengthens emotions in the body, leading to varying levels of neurosis. Representing thoughts in the form of words is a valid approach in this case for a number of reasons. One is that most people would say that a

large part of their thought processes, particularly higher level reflective thinking, occurs in the form of words in their native language. Second, through the development of verbal abilities in humans over thousands of years, language has become almost inextricably intertwined with the realm of thought and has spawned the great meta-domain of psycholinguistics. A great deal of research has been conducted in the relationship between thought and language. Third, from a technical standpoint, using words in an interface allows for great deal more flexibility and variability in user input as well as computer analysis than does using images or some other form of visual symbolism.

### **3.3 Thought-Formation Mechanics**

A variety of mechanics for inputting, or “forming” thoughts in the experience were tested. Gradually it became more apparent that the system more open and generative, but while preserving the psychological integrity of the simulation. For instance, one of the principles originally imbedded in the system was that a positive thought would more likely be followed by another positive thought than a negative one, and vice versa. Having a word pool from which users could create thoughts would allow the system to change the word pool according to every thought that was formed, increasing the relative numbers of one type of word and decreasing the number of other types. This manipulation was based on the premise that users would be looking to make things “happen” in the experience and would not interact with the system in a way that is truly reflective of their thought patterns. It was also based on a desire to convey to users the notion that their thought processes are *not* as free as they might think. Under actual life circumstances, certain personal habits may impede positive thinking or subconscious agents can take over. Moreover, as already mentioned in Chap. I with reference to Carl Jung, thoughts are not as unique to each individual as it may seem. We are limited by certain universal predispositions.

#### *3.3.1 Prototyping Technology*

Based on the metaphor of the collective unconscious, the application was designed as a simultaneous multiple user experience – many users affecting one avatar with their thoughts. However,

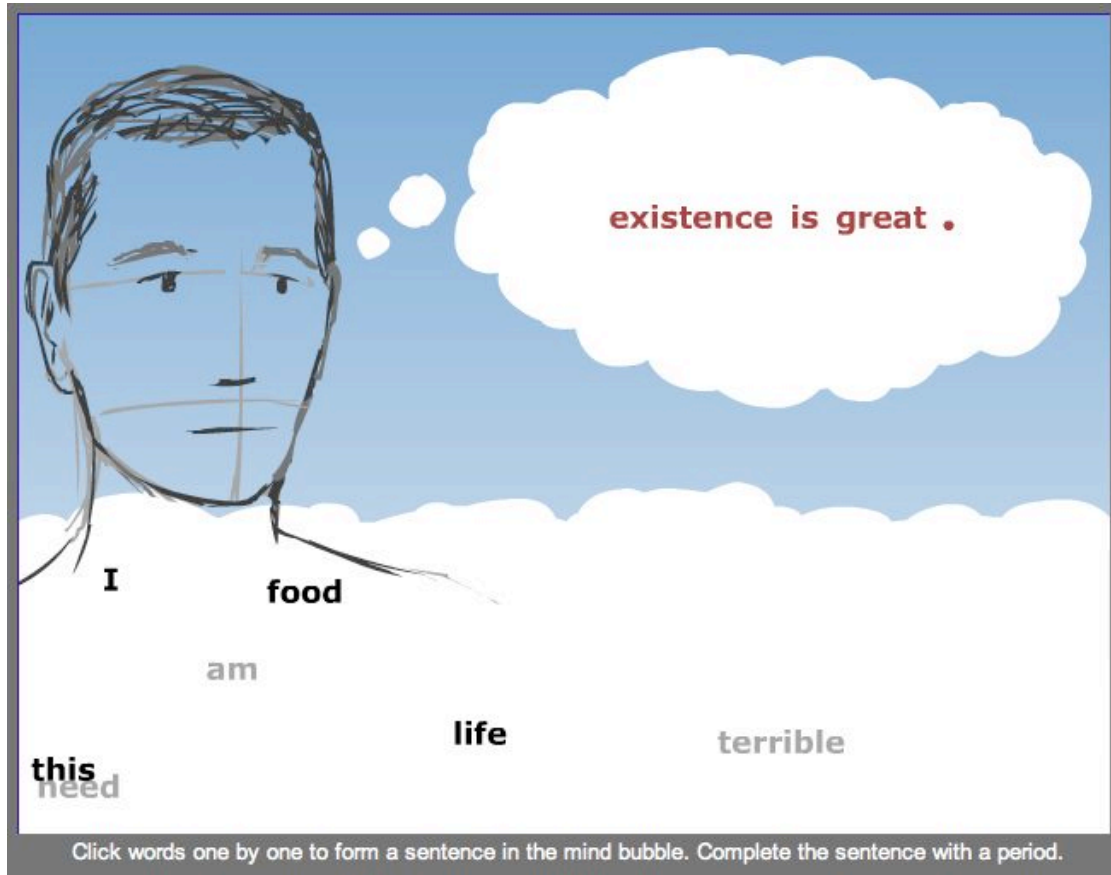
before dealing with the complexities of involving multiple users, and in order to make the system as entertaining for a single user as it would be for many at once, this author sought to quickly simulate a single user's interaction. To prototype the thought-formation mechanic, Macromedia Flash was chosen as an authoring tool. Its native programming language, Actionscript, is ideally suited for developing richly interactive and dynamic online multi-user experiences like this one.

### 3.3.2 *Active Thought Formation: Dragging and Clicking on Words*

In the first prototype devised, the user and the computer take turns dragging a word or phrase from a moving stream of words into the universal man's (avatar's) "mind bubble". The computer checks the set of words currently in the bubble against a list of meaningful sentences (stored in a database) and as soon as it finds a match, it quickly displays the words in the appropriate arrangement to form the matching sentence. Afterward, the next word thrown into the bubble clears the sentence and starts a new thought. At this early stage, there was a possibility that the experience could be a game in which users would try to produce positive thoughts while avoiding negative ones. The involvement of the computer in this case was intended to introduce a bit of a challenge by making it more difficult for users to ensure that a positive thought would form in the mind bubble. In testing the mechanic, seeing the computer suddenly shoot a random word into the mental mix proved to be too disruptive and disconcerting.

In addition to the above, it became apparent from analyzing this prototype that users would want to form a thought more quickly and be given more flexibility in the way that they combine words. A later version eliminates the set of prescribed sentences and allows any combination of words to be used to form a sentence up to a two-line limit. To encourage users to form grammatically sound verse, the set of test words were categorized by part of speech and a simple set of grammar rules were written in an XML document for Actionscript to parse and check for basic noun-adjective-verb agreement. Words that grammatically follow previous ones in a sentence are highlighted in the word stream while those that do not are faded, providing a word-by-word guide for users to follow in forming sentences. The program also contains rules to detect when a complete sentence had been formed so that the mind bubble can be

cleared for the next thought. In addition, the mouse mechanic was converted from “drag-and-drop” to a simple click, which feels less tedious and more intuitive.



**Fig. 1** Sketch prototype of active thought formation based on grammar rules

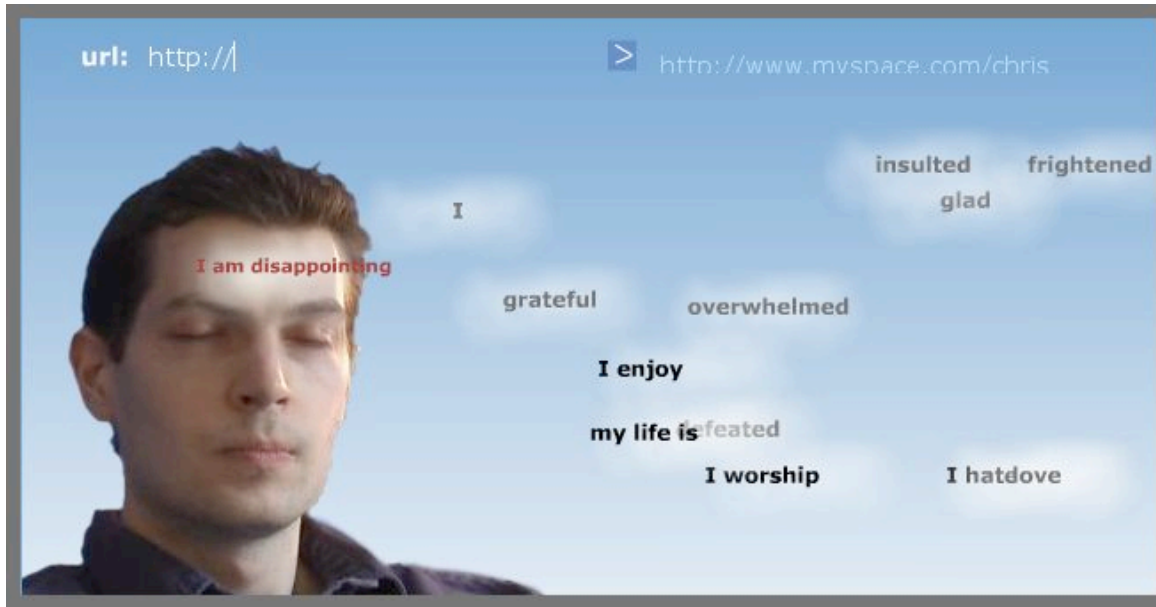
User testing of this prototype was premature in that it did not yet incorporate code to bring in new sets of words after each thought was formed, but nonetheless this author took into consideration the sense of restrictiveness and repetitiveness that users felt when forming thoughts. The next step involved thinking about ways in which the system’s vocabulary could grow on its own. Using a dictionary or thesaurus in web-based format would potentially allow the program to add new words dynamically to the database by assigning them the same emotional associations as their synonyms. With this method, it would be possible to allow users to contribute thoughts in a free-writing form prior to run-time or to extract content from blogs on a regular basis. In one scenario, this author envisioned users clicking

through a thought formation guide based on constantly evolving collections of sentences from free-writing forms or Internet weblogs. Testing this mechanic in Flash revealed its many flaws, the primary one being that the user often becomes locked into finishing a sentence that he does not want to finish. In addition, treating written sentences as being indicative of natural thought patterns, even when coming from a so-called “stream-of-consciousness” blog, is problematic.

### 3.3.3 *Automated Thought Formation*

It became clear that designing this type of experience to be approachable, intuitive and engaging would be very difficult based on its philosophical underpinnings. When given free reign to interact with a system, the user’s primary instinct is often to accomplish a goal or make something very tangible happen rather than to engage in a mental discourse with the interconnected elements of a system. Hence a pivotal step in the process was deciding to significantly alter the role of the user. The experience was changed from an immersive “game” in which the user is actively involved, to a largely automated spectacle. In this way, users are encouraged to take more of an observational stance rather than becoming frustrated with the results of their actions.

In the subsequent iteration of the project, the words intelligently combine into thoughts by themselves based on a set of grammar rules. Once a “full sentence” forms, it enters the avatar’s head. The user interacts simply by typing the URL of a web page, such as a blog, into a text field. The system extracts the emotional words from the web page, along with some random pronouns, auxiliaries, and universally relatable nouns such as job, family, and money, and places them into a chaotic mix where they quickly combine with other words whose part of speech is in accordance with their own. A pleasing aspect of the experience is that neither the user nor the system’s author has direct control over the outcome. The user intuitively adds URLs that he believes will contribute to a certain emotional response by the avatar, and derives pleasure from the confirmation of his expectations or from a surprise outcome.



**Fig. 2** Prototype for a semi-automated stream of consciousness virtual experience

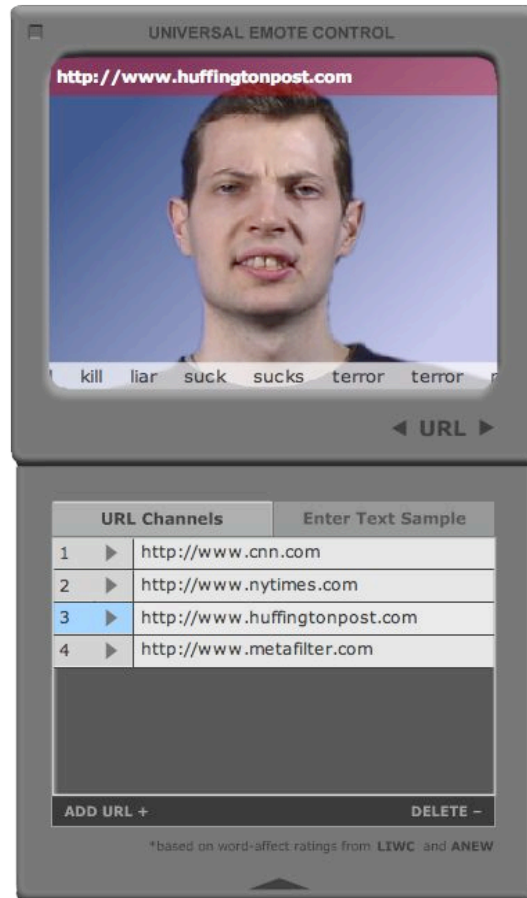
### 3.3.4 *The Final Design: Words without Thoughts*

Despite its significant advances from previous thought-formation prototypes, this author decided that the words should take on a less overt and externalized role. The final prototype, the Universal Emote Control, eliminates actual thought-formation from the application for several reasons. One is that the thoughts tend to produce very predictable and repetitive results on the avatar's face. The user does not get an *overall* sense of a web site's emotional effect on the avatar, and in many cases the expressions hover around neutral for long periods of time. Another reason is that designating emotions for single sentences based on keywords is problematic when context is not taken into account. (See 3.5 *Detecting Affect in Text through Psycholinguistic Analysis*). Finally, developing flexible, but efficient grammar rules can be a difficult and unending task.

As described briefly in Chap. I, the Universal Emote Control is a customizable "widget" — a micro-application that users can download to their desktops. It automatically connects to web pages that users specify in the form of URLs such as [www.cnn.com](http://www.cnn.com), the address of their own blog, or any other site from which text can be accessed by a server-side script. The interface is designed to evoke a television set with a remote control attached. The avatar's face appears on a "screen" framed at the top and bottom by

two bands, one showing the current URL selected, and the other functioning as a ticker of emotion words extracted from the web page. Users can open or close a remote control tray in which they create and manage their own list of URL “channels” whose content the avatar processes emotionally. With the URL forward and back buttons located directly beneath the screen, the user can skip to the next or previous URL on the URL channel playlist. At regular intervals, the widget automatically checks for updated content on the web pages listed in the tray. In another panel, the user can paste some text, such as a piece of creative writing, into a text field and have the avatar process it emotionally.

The notion that the avatar is processing the text of the web pages mentally is now much more implicit. Though it does not provide the same breadth of insights into the nature and mechanics of human thought as was originally intended, the design effectively focuses the user’s attention on a basic input and output effect and is more revealing in a way that ordinary Web users can relate to.



**Fig. 3** Interface of the final working prototype of the 'Universal Emote Control.'

### 3.4 The Word Database

Driving the interface of the project is a large word database created with Microsoft Access. In order to conduct word searches on text samples and to display words on screen, the Flash application connects to the database through Macromedia Coldfusion, a popular middleware language for Internet-based applications.

### 3.5 Detecting Affect in Text through Psycholinguistic Analysis

The project draws extensively from two psycholinguistic data resources, L.I.W.C. and A.N.E.W., introduced in Chap II. They are well regarded as tools for analyzing emotional affect in text. L.I.W.C. was specifically designed to employ a "word count" strategy in text analysis. (Pennebaker, 4) Word count

programs are algorithmically simple and efficient in that they make psychological inferences based on cumulative keyword detection. (Pennebaker, 8) However, it is important to note that this keyword-based method is most statistically meaningful for large bodies of text as opposed to short ones. Word count strategies cannot account for irony, sarcasm, or the context in which a word is employed. (Pennebaker, 8) For instance, making the assessment that a single sentence such as “I have beaten the odds” is inflected with anger because it contains the word “beaten” would be considered circumspect. However, if the word “beaten”, or semantically similar words, appear multiple times throughout a piece of text, it is more likely that a subtext of anger is involved.

### 3.5.1 *L.I.W.C.*

Linguistics Information and Word Count designates 615 words as being indicative of one of five affective processes – positive feelings, optimism and energy, anxiety/ fear, anger, and sadness/ depression. Based on these categories, the avatar system is modeled on four primary emotions – *happiness/ positive feelings, anxiety/ fear, anger, and sadness/ depression*. (Optimism and energy words were dropped). The desired word sets were parsed from individual text files into the database and mapped to their designated emotions.

### 3.5.2 *A.N.E.W.*

In order to significantly broaden the scope of words recognized by the system and to add another level of emotion data to the mix, words from A.N.E.W. were used to supplement the LIWC data. A.N.E.W. contains a large set of words, many of them nouns, which are not typical emotion words, but are nonetheless emotionally inflected. About a quarter of them also appear in L.I.W.C., but instead of being categorized by emotion, they are rated along three primary dimensions — pleasure, arousal, and dominance — by both male and female “judges”. Each dimension is represented as a rating scale from 1.00 to 9.00. Pleasure, or valence, is a measure of how happy vs. unhappy the word makes the subject feel (5.00 indicates neutrality). Arousal is a measure of how aroused/ excited vs. calm the word makes the

subject feel, whether it is positive or negative. Dominance is a measure of how controlled vs. in-control the word makes the subject feel.

Making use of the A.N.E.W. words in conjunction with L.I.W.C. words required a method of translating pleasure, arousal, and dominance values into the four primary emotions designated above. This author devoted considerable time to analyzing the ratings, looking for discernable patterns in a sampling of the words, and came up with the following deductions:

**“HAPPY”** words: pleasure > 6.75

ex: kindness, excellence, hopeful, holiday, joke, miracle, love, interest, acceptance, cute

**“FEAR”** words: pleasure < 4.75, arousal > 5.50, dominance < 5.00

ex: afraid, ambulance, bankrupt, cancer, demon, fearful, horror, agony, danger, guillotine, insane, jealousy, leprosy, lost, killer, knife, insult, abuse, dead, damage

**“DEPRESSED/SAD”** words: pleasure < 4.75, arousal < 5.50, dominance < 5.00

ex: cemetery, bored, death, deformed, disappoint, dreary, coward, discouraged, illness, ignorance, failure, impotent, idiot, gripe, handicap, loser, morbid, misery, meek, lonely, bereavement, abortion

**“ANGRY”** words: pleasure < 4.75, arousal > 5.00, dominance > 3.50

ex: anger, evil, demon, gripe, guillotine, insane, jealousy, leprosy, morbid, misery, mad, killer, knife, insult, horror, agony, abortion, abuse, damage

One will notice that certain words have been placed in more than one category, as some of the words connote mixed emotions (where arousal and dominance ranges overlap). In order to validate these qualifying numbers, this author enlisted the aid of a colleague in a Psychology and Design course. The colleague, a PhD candidate in psycholinguistic theory, used T-Tests in a program called Statistical Package for the Social Sciences (SPSS) to analyze the word groupings and their individual values in order to determine if the differences between each pair of word groups was statistically significant. The results

showed that they were in fact significant, as p values came out lower than .05 for all pair-wise comparisons. The most nebulous and least refined distinction, nevertheless remains between fear and anger words, as they only differ in one dimension (dominance), the degree to which the subject feels in control. However, many words that connote fear also connote anger, and hence the dominance values overlap significantly.

**"FEAR"** and **"DEPRESSED/SAD"** significantly differ on the following dimensions:

Valence:  $t(40) = -2.168$ ;  $p = .036$

Arousal:  $t(40) = 10.177$ ;  $p < .001$

**"FEAR"** and **"ANGRY"** significantly differ on the following dimension only:

Dominance:  $t(37) = -2.424$ ;  $p = .020$

**"FEAR"** and **"HAPPY"** significantly differ on the following dimensions:

Valence:  $t(28) = -26.421$ ;  $p < .001$

Dominance:  $t(28) = -8.909$ ;  $p < .001$

**"DEPRESSED/SAD"** and **"ANGRY"** significantly differ on the following dimensions:

Arousal:  $t(39) = -7.800$ ;  $p < .001$

Dominance:  $t(39) = -3.230$ ;  $p = .003$

**"DEPRESSED/SAD"** and **"HAPPY"** significantly differ on the following dimensions:

Valence:  $t(30) = -20.142$ ;  $p < .001$

Arousal:  $t(30) = -5.374$ ;  $p < .001$

Dominance:  $t(30) = -9.059$ ;  $p < .001$

**"ANGRY"** and **"HAPPY"** significantly differ on the following dimensions:

Valence:  $t(27) = -24.287$ ;  $p < .001$

Dominance:  $t(27) = -5.751$ ;  $p < .001$

Though refinements could in fact be made to the qualifying ranges, the categorizations were deemed significant enough to be applied as a general rule across the entire A.N.E.W. word list. All words that fall within the categories set by this author were parsed into the database and automatically assigned an emotion id as they were for L.I.W.C. words. The majority of the emotion ids for words listed in both A.N.E.W. and L.I.W.C. were consistent with one another, but where there was a discrepancy, L.I.W.C. emotion ids took precedence. In addition to being used to determine the emotion ids for each word,

pleasure and arousal values take on important individual roles in calculating the avatar’s facial expressions. (See 3.8 *Mapping Word Ratings and Designations to Facial Expressions*)

word_id	word	emotion_id	valence	arousal	dominance	
77732679	<a href="#">bland</a>	15	4.1	3.29	4.88	[ <a href="#">delete</a> ]
77732666	<a href="#">blasphemy</a>	15	3.75	4.93	4.75	[ <a href="#">delete</a> ]
77732294	<a href="#">bless</a>	14	7.19	4.05	5.52	[ <a href="#">delete</a> ]
77732552	<a href="#">blind</a>	15	3.05	4.39	3.28	[ <a href="#">delete</a> ]
77732203	<a href="#">bliss</a>	14	6.95	4.41	6.12	[ <a href="#">delete</a> ]
77732589	<a href="#">blister</a>	15	2.88	4.1	3.98	[ <a href="#">delete</a> ]
77732425	<a href="#">bloody</a>	17	2.9	6.41	3.96	[ <a href="#">delete</a> ]
77732146	<a href="#">blossom</a>	14	7.26	5.03	5.53	[ <a href="#">delete</a> ]
77732582	<a href="#">blubber</a>	15	3.52	4.57	3.86	[ <a href="#">delete</a> ]
77732243	<a href="#">blue</a>	14	6.76	4.31	5.63	[ <a href="#">delete</a> ]
77732305	<a href="#">bold</a>	14	6.8	5.6	6.67	[ <a href="#">delete</a> ]
77732767	<a href="#">bomb</a>	16	2.1	7.15	4.54	[ <a href="#">delete</a> ]
77732369	<a href="#">bomb</a>	17	2.1	7.15	4.54	[ <a href="#">delete</a> ]
77731902	<a href="#">BORE</a>	15	2.95	2.83	4.11	[ <a href="#">delete</a> ]
77731903	<a href="#">BORING</a>	15	2.95	2.83	4.11	[ <a href="#">delete</a> ]

**Fig. 4** Sections of the “Words” table, in alphabetical order, showing emotion words and their corresponding values

emotion_id	emotion
14	<a href="#">happiness</a>
15	<a href="#">sadness/ depression</a>
16	<a href="#">fear/ anxiety</a>
17	<a href="#">anger</a>

**Fig. 5** The “Emotions” table showing the major emotions recognized by L.I.W.C.

### 3.6 Connecting to Web Pages

In order to access the emotional content of a specified web page, this author wrote a Coldfusion script to search through the page's source file, strip away all extraneous tags and characters, including Javascript, styles, HTML, and comments, and scan the remaining words for any and all matching emotion words, including duplicate instances, in the database. For each matching word, the script retrieves its corresponding emotion id and ratings and sends them to Flash. The set of methods by which Flash connects to a Coldfusion or other server-side script to retrieve these values is referred to as Flash Remoting. The idea of analyzing a web page is a good one for text analysis since web pages typically involve large quantities of text. The only downside to the technology is that the text of some multimedia sites, such as ones created in Macromedia Flash, is inaccessible. Users are thus limited to entering URLs of HTML-based pages.

### **3.7 Avatar's Facial Expressions**

The face of a "universal man" serves as a relatable and whimsical means through which the emotional content of text can be expressed. From the beginning, this author sought to create a model of human facial expression that is both visually realistic and structurally modular, allowing for emotions to mix and gradate in unpredictable ways.

#### *3.7.1 A Simple Test in Flash*

Before shooting video, this author created a simple sketch rendering of the avatar's head in Flash and isolated the mouth and eyebrows of the sketch as separate movie clips so that they could be programmatically adjusted independently of one another to form a variety of facial expressions. (See *Fig. 1*) Rotating the corners of the mouth downward, for instance, produces a frown, while rotating them upward produces a smile. When tested, users enjoyed seeing the various combinations of anger, joy and sadness that could evolve from the automated shifting (in position and angle) of the eyebrows and corners of the mouth.

### 3.7.2 *A Modular Approach to Interactive Video*

This author was able to achieve a similar degree of modularity and emergence to that described above, with video-based facial expressions. The process involves videotaping three different sections of the face performing separate muscular movements and blending them together seamlessly within the application. Each facial component (eyebrows/ forehead, eyes, and mouth/ lower face) is processed as a separate semi-transparent image and given separate instructions based on the particular emotion being processed. In devising the facial expressions, this author referred to Paul Ekman's *Emotions Revealed*, a book that shows numerous photographic examples of universal facial expressions and what they signify. A multitude of different types of emotions can be conveyed primarily through subtle shifts in the relative positions of the eyebrows, eyes, and mouth.



**Fig. 6** A modular approach to interactive video based on combining image sequences of three separate facial components to form a variety of expressions

### 3.7.3 *Developing the First Prototype*

This author's own face served as the model for the first prototype. The process involved framing the head against a hand-made blue screen and taping a smooth progression of mouth/lower face movements (from smiling joyfully to mouth neutral to lips pursed to mouth agape), a progression of

eyebrow movements (from brows knit to brows neutral to brows raised to brows angled upwards), and a simple eyelid progression (wide open to fully closed). It was important to keep the head still for all of the takes so that all three components would stay consistently aligned with one another when assembled together in the computer. iMovie was used to capture the video footage and export the segments to Adobe After Effects. With After Effects, this author applied a chroma key effect to eliminate the blue background in each clip, applied a feathered mask to visually isolate the particular area of the face with which the clip was associated, and created a composite of the layers to tweak the blending and test out the animation. The final step involved exporting each clip separately as a PNG image sequence, reducing the frame rate by an average of about 60 percent to minimize the total number of images. PNG is the most optimal image compression format that supports alpha channel blending, but the file sizes (40K each) nonetheless take a large toll on Flash's processor when many images are involved. (Video formats such as QuickTime interpolate the frames and yield a much more compact file size, but do not appear to support the necessary alpha channel blends). The final prototype remedies this problem by loading the PNG images from an external directory. (See 3.6.4 *Refining the Representation of the Avatar*)

In Flash, this author imported the image sequences into separate timelines and wrote a script to move backward or forward along each one, frame by frame, a particular number of frames at a time, depending on the emotion and intensity value of the particular word being processed. This author developed a simple prototype in which a variety of emotion words stream behind the avatar's head and the user can click on a word to make it fly into the forehead and produce a reaction on the face. Each emotion cumulatively builds upon the previous one. An arsenal of angry words causes the avatar's brows to become increasingly knit up to a certain limit. A combination of happy words and fear-inducing words produces a curious expression in which the avatar appears to be masking his fear with a smile.



**Fig. 7** A prototype test for facial expressions

Despite some of its technical successes, the test revealed that further work needed to be done in enhancing the variability and range of the facial expressions. However, “branching” to new video clips within each facial component seems to contradict or work against the idea of making video modular. It was an option that was only considered for a brief period of time.

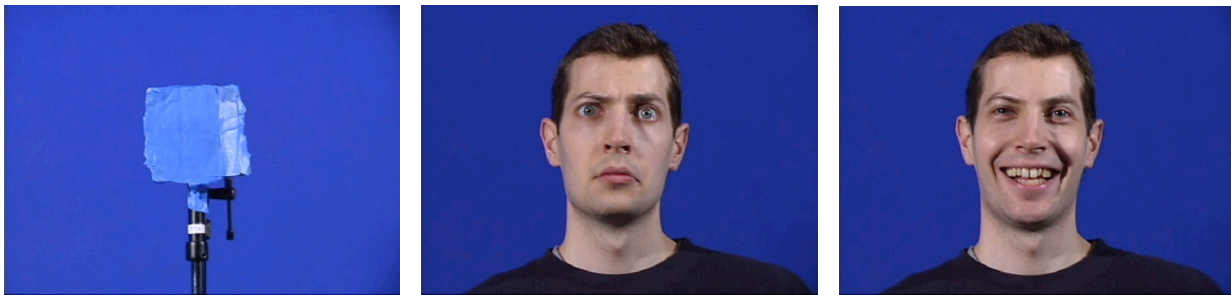
#### 3.7.4 *Refining the Representation of the Avatar*

In a subsequent iteration, this author enlisted an older male actor to play the role of the “universal man”, looking to remedy some of the calculated and robotic feelings engendered by the first performance and to increase the overall entertainment value. The actor was taped in a blue screen studio with a full lighting kit and XL1-S Pro mini-DV video camera, and a headrest apparatus was prepared to assist him in keeping his head still. This author gave the actor specific instructions in conveying the four primary emotions and in addition to capturing several takes for each of one of them, allowed time for a number of miscellaneous takes of spontaneous expressions that could potentially be incorporated into the mix. However, it became clear from analyzing the resulting footage that there were a number of technical and

conceptual issues at hand. Aside from minor mistakes in the lighting, focus, and aspect ratio, the actor's very ordinary and average-Joe-like appearance made him unconvincing as an exhibitor of emotions. In particular, the physiognomy of the actor's face prevented him from stretching classical facial expressions to their limits.



**Fig. 8** Stills from video footage of the actor initially tagged to play the role of “universal man”



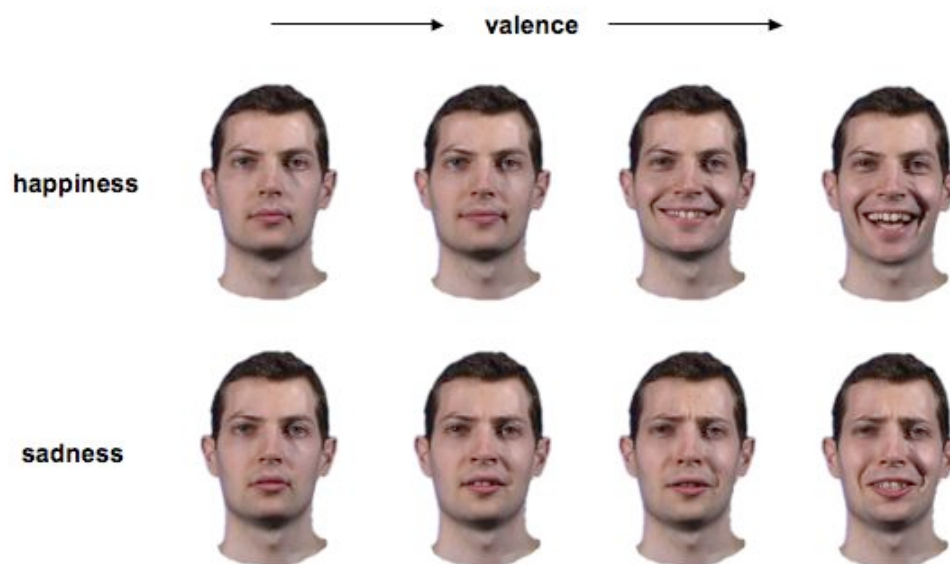
**Fig. 9** Stills from the final round of video footage showing the headrest and the author as the avatar

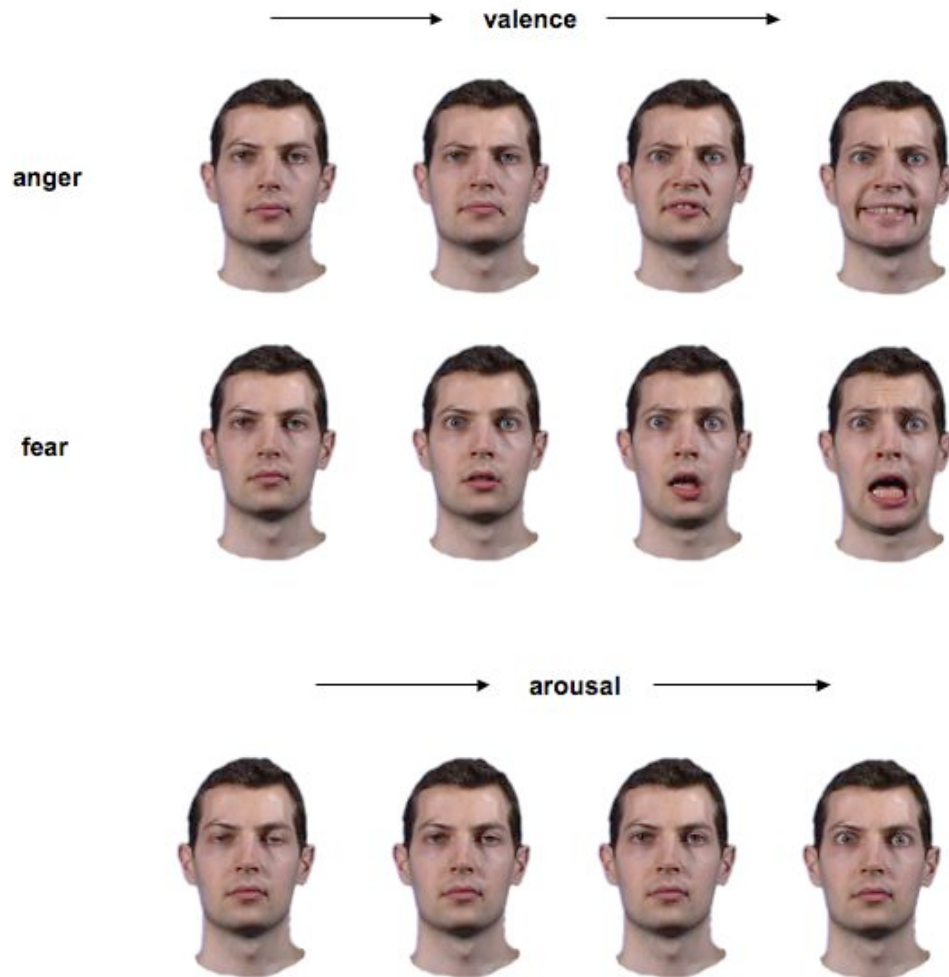
Deciding to abandon the actor's footage and return to this author's own visage for the final prototype proved to be a wise decision in retrospect. Re-shooting the video footage allowed this author to correct the mistakes made during the prior shoot and to come up with a more targeted and effectively exaggerated performance, one that exploited the inherently flexible features of the subject's face. Comical and exaggerated expressions, as opposed to subtle and nuanced ones, serve to heighten the overall impact of the product. This author decided to give the avatar a unique name in line with the brand, one that was simultaneously individual and universal, as well as humorous — *EmoJoe*.

In the final prototype of the Universal Emote Control, timeline-based programming and animations are eliminated, except in the implementation of the eyes component (the avatar blinking). Instead, static images are loaded from remote directories outside of the Flash movie. There are nine image sequences altogether, each contained within its own directory: *brows\_sad*, *brows\_fear*, *brows\_anger*, *mouth\_happy*, *mouth\_sad*, *mouth\_fear*, *mouth\_anger*, and *eyes*. *happy\_brows* was eliminated because the eyebrows and forehead do not make a significant contribution in expressing happiness on the avatar's face. When each PNG image is set to load in the right position over the background image of the avatar's face, the blending is automatically seamless, as it is in Adobe After Effects.

### 3.8 Mapping Word Ratings and Designations to Facial Expressions

The application maps emotions, arousal, and valence values to facial expressions by making a number of procedural calculations in Actionscript. Valence values of words collectively determine the extremity of a particular emotion, as conveyed by the eyebrows/ forehead and mouth/ lower face components. Arousal values collectively determine the degree to which the avatar's eyes are open vs. closed.





**Fig. 10** Sample images from the full 40-100 images in each original emotion sequence

For any text sample or URL entered by the user in the Universal Emote Control, the avatar's facial expression is calculated as follows:

1. Scan text and return all matching emotion words from database, including duplicate instances
2. Get associated *emotion id(s)* for each matching emotion word (14, 15, 16, 17)
3. Calculate *total valence* for each set of emotion words by id and determine the *highest* and *second highest valence*

4. If the *total happy (positive) valence* < *total valence* for all negative words, make the *second highest valence* the *highest valence* and make *total happy valence* the *second highest valence*
5. If *highest valence* > *second highest valence* + (big number), just use the emotion with the highest valence and drop the *second highest valence*
6. Emotion with *highest valence* becomes *brows-forehead* emotion  
Emotion with *second highest valence* becomes *mouth-lower face* emotion  
(do reverse if *total happy valence* = *highest valence* because mouth is primary indicator of happiness)
7. Get emotional density of the text (*total # emotion words / total word count*)
8. *highest valence* \* *emotional density* \* modifier = image sequence no. for *brows-forehead* emotion
9. *second highest valence* \* *emotional density* \* modifier = image sequence no. for *mouth-lower face* emotion
10. Calculate *total arousal* for all words of a certain valence (at least 2.00 points above or below 5.00 to inhibit *average arousal* from evening out too much). Calculate *average arousal* from these words
11. (*average arousal* \* eye image sequence length) / *maximum valence* [9.00] = image sequence no. for eyes

### 3.9 Developing for the Desktop

Deploying the widget to the desktop for the Macintosh platform involved investigating a number of third-party SWF-to-desktop plug-ins for Flash as the Flash application does not natively support custom-shaped, “windowless” projectors or projectors that change shape. In the case of the Universal Emote Control, the widget’s shape changes when the user opens or closes the drawer containing the list of URLs. mProjector and MDM Zinc are popular commercial options, though because of performance issues and poor support for certain features such as masking, this author instead decided to use a trial version of a more independent product called SWFCargo. SWFCargo’s technical support proved to be outstanding. Not only did they resolve certain issues that came up with regard to desktop display and the SWFCargo API, but they were quick to add new features, such as Flash Remoting, at this author’s request.

### 3.10 User Testing and Refinements

A formal user testing session was conducted with five participants to assess the usability and efficacy of the Universal Emote Control. Each participant was asked to enter specific URLs in the channel list, to paste a sample short story into the Text Sample panel, and to write down the emotion(s) they were seeing on EmoJoe's face for each. They were also asked to evaluate how "useful" they found the product and to rate certain proposed enhancements, such as animation in the avatar's face, up-to-the-minute emotional updates, an "Average All URLs" function, and the ability to customize EmoJoe's appearance.

The results of the survey were quite varied. At least two users misread a slightly happy expression as neutral. In addition, users chose very different words to identify mixed facial expressions or were unable to name the expression. However, this inability to easily label an emotion may not be a defect of the product so long as the expression resonates in some way with the site, the text, and/ or the emotion words being displayed. None of the participants objected to the expressions they saw based on the URLs entered, and most found them appropriate in some way. The widget's efficacy as an entertainment device seemed to prevail over notions of it being useful, though one user commented that seeing the face change based on the news would be great. In terms of enhancements, users preferred different things. Most popular was the ability to customize EmoJoe's appearance, and least popular was a function that calculates an average expression based on all URLs in the user's channel list.

Based on the results, this author decided to increase the facial expression modifier to elevate the occurrence of emotional extremities. Subtle and nearly neutral expressions do not appear to captivate users. In addition, visual plans for an EmoJoe customization panel were diagrammed (see *4.4 Future Directions*), eye blinking was added, and the loading scheme was revised so that users can revisit previously viewed URLs without having to wait for them to reload. The application was also programmed to automatically check for updates in each URL every 50 seconds and to deliver a new expression to the current URL if the emotion words change. This feature is one that requires further development and testing. Creating smooth transitions from one expression to another is difficult, if not impractical, to

achieve when loading images externally, and few news sites and blogs change frequently and dramatically enough within a short amount of time to make the updates evident to the user. One solution that has already been employed involves placing statistical emphasis on headlines by searching for the first item delimited by a header tag in the URL's source code. The program multiplies the valence of any emotion words that appear in the headline by a modifier to elevate its effect on the total valence for the associated emotion.

## 4: Evaluation

### 4.1 Summary of Project Outcome

The consummation of this thesis has included the development of a high fidelity prototype for one of several key web-based applications comprising the *Universal Emote* brand. The brand evolved from honing the central idea of an avatar's emotional relationship to words and extending it to the World Wide Web. In its final instantiation, the Universal Emote Control prototype appears to be a success in its execution, clarity, and usability. With more production time, it would be possible to improve the efficacy of many of its features to make the product more compelling to users over extended periods of time.

#### 4.1.1 *Concept and Meaning*

The development of the project has been guided by one central concept: collective, universal thought patterns being channeled into the mental-emotional body of a single human being represented as an avatar. In the project's evolution from a virtual experience to a whimsical commercial application, this author strove to maintain a certain conveyance of philosophical meaning — spiritual and psychological metaphors about the interrelationship between thoughts and emotions, integrating the wisdom of Jungian psychology, the Law of Attraction, and Zen. However, the simplification of form that the project underwent in order to attain greater focus may have compromised some of its depth if not its originality. The removal of poetic, narrative and expressive elements, including the representation of the avatar's life, as well as the increasing emphasis on hard statistical data may have left the product in a more superficial and mechanical state. In exchange however, it inherited a heightened level of commercial appeal. A number of individuals who have seen or heard about the project's concept have pointed to its potential as a marketing tool.

The final prototype presents some ambiguity in its lack of distinction between thought and language and in the uncertain identity of the avatar. Are the words that stream by on screen indicative of

the avatar's thoughts or are they external stimuli? Perhaps they are something in between, as language is the link between our minds and the world. Who is the avatar? Is he universal? Or is he an unwitting individual bombarded by emotions from cyberspace? Is he a typical human? Hypersensitive? Or does persona not matter? These questions are not given a satisfactory answer.

#### 4.1.2 *Design Components and User Experience*

The Universal Emote Control is visually designed to invoke, or perhaps heighten, an association of the World Wide Web with the television medium. Aside from the obvious parallel to a “universal remote”, it likens a computer's window and its buttons to a television set and remote control, with a “live” video and information feed, URLs analogous to channels, and an on-screen ticker characteristic of cable news networks. The whimsical comparison to TV not only helps to make the product approachable and amusing, but offers a sense of physical presence and expression within a virtual universe overrun by disembodied pages of sprawling text. It lends a humanistic component to cyberspace — a component that has been intrinsic to the television medium since its inception.

Because of its use of video-based images, some users may expect more movement and expression from the avatar. While this author worked diligently to introduce a greater illusion of life and spontaneity in the avatar, this thesis has led to the realization that there is a trade-off involved in designing facial expressions that are driven dynamically by code and which adhere faithfully to data. A high degree of control, including keeping the avatar's head still, is necessary to make the automated assembly of emotions possible. This is especially true when trying to maintain a realistic appearance in the avatar.

The use of video as opposed to animation to portray the avatar has both its benefits and its limitations. Its advantages include *a better fit for the overall concept* (an anonymous individual being affected by cyberspace rather than representing it), *refreshing comic appeal* (users are very accustomed to seeing cartoon avatars and would not be as surprised to see a peculiar unabated expression on that type of avatar), and a *relatively short production time*. Its disadvantages include *a low degree of malleability and real-time customization capabilities* (the geometry of the avatar's face is not accessible as it is in a 3D

model and hence it is not easy to make changes to the avatar's appearance) and *larger file sizes* (particularly for sequences of images with alpha channels). The file size issue does not carry significant weight, however, when the images are loaded externally (see *4.1.2 Methodology and Process*). Despite the first drawback, it is worth noting that the majority of 3D avatar customization modules do not offer specific facial feature options. Users can select from a number of predefined templates or physical character models and then personalize certain attributes such as hairdo, hair color and clothing but usually cannot adjust the actual physiognomy of the face.

Currently, the facial expressions exhibited by the avatar do not convey maximum resonance with the URLs being displayed, particularly news sites with mixed stories and mixed emotion words. Seeing positive words in the ticker when the avatar has a negative expression, even though the negative words statistically outweigh the positive ones in valence, is problematic. In this sense, the avatar's expression bears the most resonance and impact for small, themed sites such as [www.suicide.com](http://www.suicide.com). Placing emphasis on headlines can be a good solution for news sites and blogs, but only if this process is made apparent to the user. Highlighting the particular emotion words found in headers to separate them from other words may not be enough

#### 4.1.2 *Methodology and Process*

The final prototype employs a unique fusion of technologies, including an online database, visual effects, dynamic text, video-based image sequences, and a graphical user interface. They have been made to work seamlessly with one another through integrative programming and design. The decision to employ particular software packages to solve particular design problems proved successful. For quick, simple, and efficient video editing with easy export to H.264 format, a quality video compression scheme, iMovie was a good choice. For efficient compositing of video clips with a wide range of export options for image sequences, Adobe After Effects was also a wise selection. For fast database creation and management, Microsoft Access was the best tool, and for effective development of media-rich Internet-based applications, Adobe's Macromedia Flash is the obvious winner. Selecting a SWF-to-desktop plug-

in for Flash required the most testing and consideration. Given the time frame, an easy-to-use commercial option with adequate speed and performance, compatibility with the Macintosh operating system, and support for the features contained in the Flash prototype, was preferable. SWFCargo proved that it can handle all of these requirements, though currently there is a processing lag in the Flash Remoting that needs to be resolved.

Extensive research into psycholinguistic analysis helped to concretize and lend greater validity to the project. A significant effect of this investigation was that it led to the recognition that the emotional meaning of a single English sentence cannot effectively or convincingly be determined by a computer, prompting the decision to eliminate thought-by-thought processing of emotions on the avatar's face. Instead, this author employed a word count strategy and based the avatar's facial expressions cumulatively on large bodies of text, incorporating word-affect ratings from two respected sources, L.I.W.C. and A.N.E.W., and devising a scheme to translate arousal, valence, and dominance values into emotional expressions. Analyzing and using the data effectively and with sound judgment is a challenge. The extent to which an artist can take liberties in the interpretation of numbers is a significant question to ponder.

Word-emotion data sources comprise an elusive area of research. The words taken from L.I.W.C. and A.N.E.W. only total about 1200 words altogether, producing a relatively low return rate on text samples. Gaining access to a more extensive source such as Wayne Chase's Connotationary or the 8700-word Dictionary of Affect in Language (D.A.L.) would help to ensure that certain affective words found in text samples would not go unrecognized. However, the downside to a high return rate (many emotion words detected within a sample) is that the processing of the text takes longer.

In developing the video avatar, the aim was not only to attain a high level of technical quality, but also to employ an out-of-the-box approach in working with the video clips. This author developed a modular system for displaying facial expressions with video, one that is potentially unprecedented. The blending of three facial components (4 video clips separated into eyebrow, eye and lower face regions) to form up to 13 different expressions with varying degrees of valence and arousal is one that successfully

reflects careful consideration of the subject at hand. Despite the seemingly great complexity and range of human facial expressions, the variety emerges largely from the coordinated movements of a few isolated muscle groups. Such a method would probably not work as effectively on a representation of the human body, which is more physically and spatially complex.

Loading the images externally into Flash by dynamically referring to their directory and identifying number proved to be a highly efficient method of displaying the facial expressions. In earlier prototypes, the images were stored and accessed inside the Flash movie itself, placing a great load on the processor. When the images are stored on a server, the number of files can be quite large without affecting the performance of the application as long as they are loaded as needed and then removed from memory.

This author's process in developing this project was a highly adaptive one. The primary intention was to remain open and flexible about the form and genre of the application from the very beginning and to explore many different types of interaction. This author refrained from giving the project a specific genre label, such as a game, as it might begin to take on the clichés of a specific domain. The various prototypes reveal that this author investigated clicking, typing and dragging text, mining words from web pages, grammar rules, automated animation, timers and other game mechanics before arriving at a user interaction model that was intuitive for the average user and at the same time, unique.

## **4.2 Future Directions**

The functional prototype of the Universal Emote Control represents a stepping stone towards more intriguing work related to the notion of an avatar as a representative of mass consciousness. Future development of this thesis involves the creation of additional word-emotion-avatar applications to entertain and enlighten the users who visit the *Universal Emote* home site.

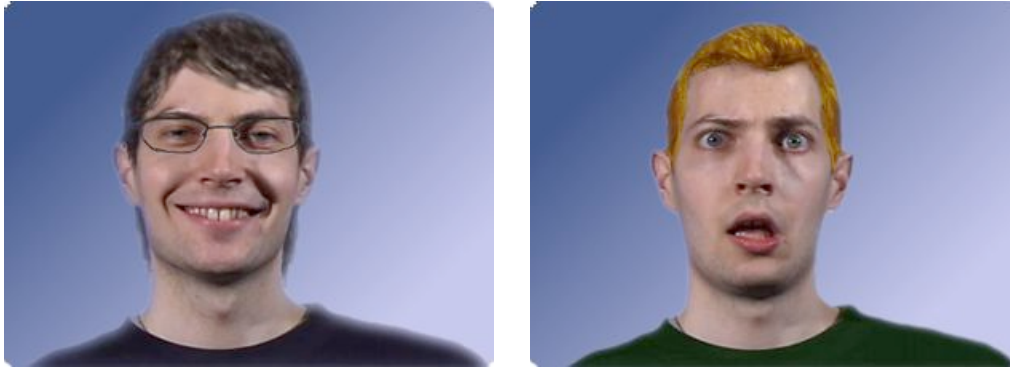
### *4.2.1 Enhancements to the Universal Emote Control*

A number of aspects of the Universal Emote Control require further development. Aside from minor technical improvements, one aspect to be considered is the specific means by which the text of URLs is analyzed and translated into facial expressions. The widget needs a method by which emotional updates can become readily apparent and as meaningful as possible. Basing the avatar's emotional memory on the dates of isolated new headlines in RSS feeds, which get updated very frequently, is one option to explore. However, making this backend analysis a bit more transparent to the user may be necessary in order for him/ her to fully appreciate the result. It may also be worth expanding the complexity of EmoJoe's facial expression model to convey more emotional information about each URL or text sample, particularly ones that are highly mixed.

Another enhancement involves giving users the power to customize EmoJoe's physical appearance, in a separate panel of the interface, thereby creating their own personal version of him. Given that it is easiest to allow dynamic modification on static features, editable aspects could conceivably include hairstyle, hair color, shirt color, skin tone, headwear and eyewear. Flash's powerful BitmapData class allows for real-time image transformations such as color changes.

#### *4.2.2 Universal Emote Control: The Online Multi-user Version*

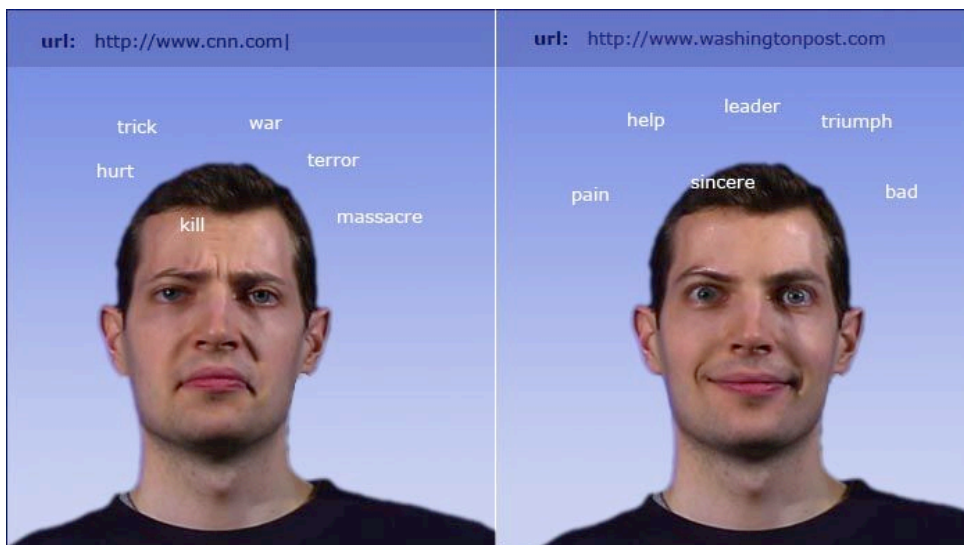
The centerpiece of the Universal Emote home site would be an in-browser version of the Universal Emote Control in which users can continually contribute to a growing master list of URLs for EmoJoe to process. The application might automatically check for link tags within the web pages entered and connect to constituent pages or external sites, aiding in the rapid consumption of the entire Internet. In another version or panel, users may be invited to contribute to EmoJoe's stream-of-consciousness in the form of a running block of text. Emotion words with a high valence and arousal would remain in EmoJoe's memory longer than those with a low valence and arousal.



**Fig. 2** Possible customization options for EmoJoe's appearance

#### 4.2.3 "Head to Head" Emoting Competition

The 'head to head' emoting competition would be a casual game in which multiple users agree on one of four emotions that they will attempt to elevate in their own version of EmoJoe. Players would quickly type in URLs in an effort to bring their version of EmoJoe to the extremity of the emotion before any of the other players. Emotion words would flow downward from the URL box and dissolve into EmoJoe's head, changing his facial expression in real time.



**Fig. 3** Preliminary visualization of the Head to Head interface

#### 4.2.4 Therapy Game

In this game, users would play the role of therapist to EmoJoe, a twist on the original purpose of an avatar. (The Lizabot, originally developed in 1966, was designed to dispense psychotherapy to its users). The task would be for user-therapists to deduce negative words infecting EmoJoe's psyche from a list of URLs currently feeding his mind in order to dissolve the associated negative emotions. Over time, users might develop a sympathetic bond with EmoJoe, and his emotional management might become critical as emotional extremities would lead to consequences of some kind.

### 4.3 Conclusion

This thesis investigation has examined the notion of a human avatar in detail and has explored a variety of techniques and ideas for representing it in unique and unconventional ways. It has looked at the avatar's use as a signifier of self and of other, its relationship to the notion of a character, how an avatar can be used in an interactive experience revolving around psychological factors, and how it can be effectively represented in video format. It has examined various contexts for avatars, including games, virtual experiences, chatterbot simulations, and other web-based applications, and has purposely sought to forge new territory, looking beyond predefined narrative structure and yet away from typical artificial intelligence simulations. In this reluctance to settle into either of these spaces, the *Universal Emote* applications represent an attempt to reconcile the apparent separation of man and machine, of mind and program, of emotion and data, of expression and code, of universe and network.

## Works Cited

“ACE Gallery | Tim Hawkinson”. 1 March 2007 <<http://www.acegallery.net/artistmenu.php?Artist=1>>.

Bradley, M.M., & Lang, P.J. (1999b). Affective norms for English words (ANEW): Stimuli, instruction manual and affective ratings. Technical report C-1. University of Florida, Gainesville, FL.

“Connotative Reference Corporation – Emotional Power Tools for Writers.” 1998-2007, Connotative.com. 1 March 2007 <<http://www.connotative.com>>.

“forget ‘emoticons’ — use your own face — tech —29 September 2006 — New Scientist Tech”, 5 May 2006 <<http://www.newscientisttech.com/article/dn10190-forget-emoticons--use-your-own-smiley-face.html>>.

Ekman, Paul. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. New York, NY: Henry Holt & Co., LLC, 2003.

Hall, Calvin S., and Vernon J. Nordby. A Primer of Jungian Psychology. New York, NY: New American Library, 1973.

Horowitz, Ken. “The Rise and Fall of Full-Motion Video.” 8 July 2005. 1 March 2007 <<http://www.sega-16.com/Rise%20and%20Fall%20of%20FMV.php>>.

“InteractiveStory.net”. 2005, Procedural Arts. 1 March 2007 <[www.interactivestory.net](http://www.interactivestory.net)>.

Johnson, Steven. Emergence: the Connected Lives of Ants, Brains, Cities, and Software. New York, NY: Scribner, 2001.

“LIWC — Linguistic Inquiry and Word Count.” 10 March 2007. <<http://www.liwc.net>>.

Pennebaker, James W., Matthias R. Mehl, and Kate G. Niederhoffer. “Psychological Aspects of Natural Language Use: Our Words, Our Selves.” Department of Psychology, University of Texas at Austin, 2003.

“Research Projects: Expressive Real-Time Communications.” 10 May 2007 <<http://www.uop.gr/~acb/projects.htm>>.

“Second Life: Your World. Your Imagination.” 2006, Linden Research, Inc. 10 November 2006. <<http://secondlife.com>>.

“The Sims”. 2005, Electronic Arts Inc. 10 November 2006. <<http://thesims.ea.com/us>>.

“TheSims2.com.” 2006, Electronic Arts Inc. 10 November 2006. <<http://thesims2.ea.com>>.

“The Sims Online.” 2003, Electronic Arts Inc. 10 November 2006. <<http://player.thesimsonline.ea.com/index.jsp>>.

Underwood, Mick. “CCMS – Communication studies, cultural Studies, media studies infobase by Mick Underwood.” 1 March 2007. <<http://www.cultsock.ndirect.co.uk/MUHome/cshtml/index.html>>.

“Vodafone.” 1 March 2007 <<http://www.vodafone.com/flash/futures/index.jsp>>.

Walczak, Marek and Martin Wattenberg. “Apartment.” 2001-2002. 5 May 2007 <<http://www.turbulence.org/Works/apartment/>>.

“Wedding Crashers Official Site.” 2005, 10 May 2007 <<http://www.weddingcrashersmovie.com/crashthistrailer>>.

Whissell, Cynthia M. “Whissell’s Dictionary of Affect in Language: Technical Manual and User’s Guide.” Laurentian University.

## Works Consulted

“DancingPaul.com: I can dance if I want to!” 1 Oct. 2006 <<http://www.dancingpaul.com>>.

Doulamis, Anastasios, Nikolaos Doulamis, Kilmis Ntalianis, and Stefanos Kollias. “An Efficient Fully Unsupervised Video Object Segmentation Scheme Using an Adaptive Neural-Network Classifier Architecture.” 2003, IEEE Transactions on Neural Networks, Vol. 14, No.3. 1 Oct. 2006 <<http://www.image.ece.ntua.gr/papers/245.pdf>>

“DRAGRI.” 26 Feb. 2006 <[http://www.dragrifan.com/en/sample\\_en/index.html](http://www.dragrifan.com/en/sample_en/index.html)>.

Egen, Sean P. “iMediaConnection: The History of Avatars.” 20 June 2005. 1 Oct. 2006 <<http://www.imediaconnection.com/content/6165.asp>>.

Hales, Chris. “Kinoautomat: The World’s First Interactive Film.” 1 Oct. 2006 <[www.kinoautomat.org](http://www.kinoautomat.org)>.

Insley, Joseph A. “VideoAvatar Library.” 1997, Electronic Visualization Laboratory, University of Illinois at Chicago. 15 Sept. 2006 <<http://www.evl.uic.edu/insley/VideoAvatars/index.html#info>>.

“Interactive Movie: Moby Games.” 1 Oct. 2006 <<http://www.mobygames.com/browse/games/interactive-movie>>.

Iuppa, Nicholas and Karl Anderson. Advanced Interactive Video: New Techniques and Applications. White Plains, NY: Knowledge Industry Publications, Inc., 1988.

Lee, Sang-Yup, Ig-Jae Kim, Sang C. Ahn, Myo-Taeg Lim, and Hyoung-Gon Kim. “Toward Immersive Telecommunication: 3D Video Avatar with Physical Interaction.” 1 Oct. 2006 <<http://vrsj.t.u-tokyo.ac.jp/ic-at/papers/2005/full131.pdf>>

Montfort, Nick. Twisty Little Passages: An Approach to Interactive Fiction. Cambridge, MA: MIT Press, 2003.

Ngan, King N., Thomas Meier, and Douglas Chai. Advanced Video Coding: Principles and Techniques. Amsterdam, the Netherlands, New York: Elsevier Science, 1999.

Salen, Katie and Eric Zimmerman. Rules of Play: Game Design Fundamentals. Cambridge, MA: The MIT Press, 2004.

“Samorost1.” 2003, Amanita Design. 1 Oct. 2006 <<http://www.samorost.net/samorost1/>>.

“Subservient Chicken.” 2004, Burger King Brands, Inc. 1 Oct. 2006  
<<http://www.subservientchicken.com>>

“The Vivid Mandala GX System: Wireless Virtual Reality Games”. 2006, Vivid Group. 1 Oct. 2006  
<<http://www.vividgroup.com>>.

Weiner, Norbet. The Human Use of Human Beings: Cybernetics and Society. Boston: DeCapo Press, 1954.

“Whissell’s Dictionary of Affect in Language for Emotional Text Recognition.” 10 May, 2007  
<[http://compling.org/~abe/emotion\\_in\\_text\\_cgi/DAL\\_app](http://compling.org/~abe/emotion_in_text_cgi/DAL_app)>.

## Appendix

### **Life Sentence.** version 2 (Full Description)

In this web-based virtual experience, users are presented with an immutable body (a shell of a person) and are invited to raise this human from a state of nothing-ness (no mind and hence no life) to a state of something-ness by giving him thoughts. As opposed to directly controlling a set of externalized behaviors, users determine the avatar's everyday routines and experiences indirectly by shaping his mind. Keywords contained within thoughts are mapped to particular emotions or to the creation of particular life structures such as home, work, the store, and school.

Users start out with a very small pool of words from which they can form thoughts in the avatar's conscious mind. A minor portion of the words are "useful" in that they can be used to form practical thoughts. Practical thoughts, which typically begin as need or desire thoughts, translate into concrete advances, acquisitions, or changes in the avatar's life, usually represented by the formation of a life structure that the avatar will subsequently visit on a regular basis. Many of the other words are negatively or positively inflected and affect the avatar emotionally while also affecting his experience inside the life structures that are formed. These experiences introduce their own specific words (negative, positive and practical) to the thought-stream, adding complexity (and potentially neurosis) to the avatar's inner life as his outer life expands.

The avatar may exhibit a variety of emotions at any time based on various combinations of words entered in thoughts. When the avatar is caught up in a negative emotion from thinking too many negative thoughts, his ability to think practically diminishes. Hence, the number of useful words available declines in favor of more useless ones. In addition, negative thoughts breed more negative thoughts and positive thoughts breed more positive thoughts. Negative emotions increase the chances of the avatar having a negative experience inside a life structure while positive ones have the opposite effect. The experiences themselves create visible short-lived emotional reactions, but only thoughts entered by users have the capacity to perpetuate and strengthen these emotions.

When the accumulation of toxic thoughts reaches a critical level (all positive thoughts have vanished from the thought stream), the avatar commits suicide, but only his mind and life evaporate. His physical body is reincarnated in a state of emptiness. According to the current plan, this outcome is inevitable in the absence of any rewards or other results, though the avatar can attain many emotional and experiential highs and lows along the way.

As a web-based online virtual experience, anywhere from one to the maximum number of users allowed by a server may enter and play simultaneously, but unlike most multi-player games, a single player alone at one time may suffice for an enjoyable user experience. The results of each user's choices during a single session are recorded by the system and have a cumulative effect (on both the avatar and his life) that future users can see.

#### *A Model of User Motivation*

Based on the principles of interaction outlined above, users will be led to adopt one or more of 3 main roles in the avatar's shared mind:

- a. The optimist: this user concentrates on seeding positive thoughts in the avatar's mind in an effort to make or keep the avatar content
- b. The pessimist: this user subversively tries to bring down the avatar emotionally by seeding negative thoughts in his mind
- c. The pragmatist/ achiever: this user tries to follow need/ desire (useful) thoughts in the midst of the chaos/ battle between optimists and pessimists in an effort to build a life for the avatar and have him be successful

These roles, of course, will often overlap and coincide with one another, as accomplishing one goal may be dependent on accomplishing another. For instance, the avatar's life success is partly determined by the quality of his thoughts. A user may also decide to give up on playing one role and concentrate on another.

None of the three paths lead to an ultimate reward. Positive feelings make the user crave more positive feelings, and desires lead to bigger desires endlessly.

LIFE SENTENCE. Basic Flowchart

