Encouraging the Expression of the Unspeakable: Influence and Agency in a Robotic Creature

by

Nicholas A. Knouf

B.S., Engineering and Applied Science
California Institute of Technology (2002)

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

Master of Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August, 2007

Author . . . .
Program in Media Arts and Sciences
20 August 2007

Certified by . . . .
Tod Machover
Professor of Music and Media
Thesis Supervisor

Accepted by . . . .
Deb Roy
Chair, Departmental Committee on Graduate Students

Encouraging the Expression of the Unspeakable: Influence and Agency in a Robotic Creature

by Nicholas A. Knouf

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning, on 20 August 2007,
in partial fulfillment of the requirements
for the degree of Master of Media Arts and Sciences

ABSTRACT

The boundary between subject and object is becoming ever-the-more blurred by the creation of new types of computational objects. Especially when these objects take the form of robotic *creatures* do we get to question the powerful impact of the object on the person. Couple this with the expression of internal, unspoken experience through the making of non-speech sounds and we have a situation that demands new thoughts and new methodologies. This thesis works through these questions via the design and study of *syngvab*, a robotic marionette that moves in response to human non-speech vocal sounds. I draw from the world of puppetry and performing objects in the creation of *syngvab* the object and its stage, showing how this old tradition is directly relevant for the development of non-anthropomorphic, non-zoomorphic robotic creatures. I show how this mongrel of an object requires different methodologies of study, pulling from actor-network theory to examine *syngvab* in a symmetric manner with the human participants. The results of a case study interaction with *syngvab* support the contention that non-speech sounds as drawn out by a robotic creature are a potent means of exploring and investigating the unspeakable.

Thesis supervisor: Tod Machover
Title: Professor of Music and Media

Thesis Reader

Thesis reader . . . .

Hod Lipson
Assistant Professor, Joint Mechanical & Aerospace
Engineering and Computing & Information Science
Cornell University

Thesis Reader

Thesis reader . . . .

Sherry Turkle
Abby Rockefeller Mauzé Professor of the Social Studies of Science and Technology
MIT Program in Science, Technology, and Society

Acknowledgments

This thesis is not an individual endeavor, much like most things in life. A number of people have helped along the way, many I have met, some I have not. I write here of my thanks with the caveat that a few bits of words are not enough to express my gratitude.

- Tod Machover, for giving me the space to explore such a mongrel of diverse interests and providing me the opportunity to develop such a bizarre project;
- Sherry Turkle, for it was in her class that I first began to consider a critical approach to the creation of technological objects. Her encouragement and support along the way have been a continual source of challenge and inspiration;
- Hod Lipson, for the inspiration I gained from his work and his pushing of the boundaries of evolutionary algorithms;
- Aden Evens, for deep discussions about the thesis and life along the way;
- All of my friends in the Hyperinstruments Group and the Media Lab as a whole, for being such wonderful late-night companions and for being there to talk about whatever strange ideas we might have;
- My professors during this journey who have always challenged me to think and rethink everything: Hiroshi Ishii, Pattie Maes, Krzyzstof Wodiczko, Chris Csikszentmihályi, Paul Horowitz;
- Linda Peterson, Pat Solakoff, and Gigi Shafer for their help in navigating the maze of the degree program;
- Jihyon Kim for being always ready to deal with my last-minute purchases and for being such a wonderful person to be around;
- The gracious support of the Things That Think and Digital Life consortia at the Media Lab;

- The developers of and contributors to numerous open-source projects. I fear that we often forget how much of our work is dependent on a community of developers whom we have never met. I especially want to thank the authors of: `Python`, `PyOpenGL`, `ODE`, `PyODE`, `liblo`, `numpy`, `scipy`, `matplotlib`, `libsndfile`, `pyaudio`, `pyaudiolab`, `PySerial`, `audacity`, `bibdesk`, `GNU/Linux`;
- Jia Yan Gu for her early work on new designs for *syngvab* ;
- John Bell, Meredith Garniss, Tatiana Fiscella, Phoebe Sengers, and David Johnson for advice, new ideas, and challenges
- Behan Pub and Matt Murphy's, and the Irish session musicians at both, for making the writing just that bit smoother;
- Diesel cafe, for making the morning writing sessions a little bit easier;
- Phil and Penny Knouf, for their love and support throughout;
- Robin Knouf, for her non-speech sounds and ever present joy;
- Orkan Telhan, for helping me on my journey into the netherlands of theory, and for many a late night talk about life;
- Monica Haller, for many, many times of intense talks about this and other projects;
- Christine Liu, a partner in many projects and always-willing and caring ear for my strange thoughts and ideas;
- Cara Frankowicz, for her humor, unwavering support, love of puppetry and wobots, patience, partnership, and friendship. Yip yip!

Table of Contents

List of Figures

List of Tables

I

Introduction: Setting the Stage

To begin by jumping into the middle of things.

This thesis is a battle against dualisms: speech–non-speech, humanities–technology, art–science, quantitative–qualitative, old–new. To fight these abstract concepts with the goal of knocking them over, of pushing them away, of banishing their thought, is bound to only create new dualisms in their place. Some say that binary thinking (coincidently, perhaps, also the thinking of computers) is inherent in the structure of our written and spoken language. Whatever the origin of these limitations in how we describe things we need to be aware that this is a war that we will not win.

So, perhaps instead I should write of an "exploration *around* dualisms". What would it mean to think of the negative space around the pillars of the old and the new? What is hidden when we focus only on the dull exterior of these ancient ideas? "Art" and "Science": the age of these words means their original brightness, their blinding light of promise, has long since faded away. We cannot claim ignorance and say, "We ignore the areas around these ideas because the ideas are too bright; they blind us and prevent us from seeing." No: a reasonable remark at some point, but not today. The intensity of these ideas is long faded. It is as if we desire something in the faded and the drab; perhaps it is comforting to think in those ways, in embracing what has lost its sheen and seems to have stood the test of years of inquiry. And perhaps we can make something anew by merely bringing art and science together, as if their combination, their *unity*, will spontaneously create a new light. An intriguing thought indeed, but the combination of two subdued colors does not make something bright. And the merging of the terms is bound to create an undifferentiated mess, a goo in which we can easily be lost, destroying what possibility might have been left deep within the ideas.

We instead must fumble our way through the dark, far away from the poles of thought. Digging around in the open space, amongst the never-touched or seen, we perhaps will find a glint of thought that is but a twinkle. With some care and polish we can see that it gives off light of its own. If we think through these new thoughts, considering their possibility but acknowledging their relationship to the dualisms that are always there, we might discover a new path anchored by a new idea. An oscillation between poles becomes a curve through space, chaotic, yes, but repeating, never. "Old" and "new", "speech" and "non-speech": they remain as attractors, with our trajectories bringing us near and far in time and thought, opening up new areas of space to explore. Had we remained anchored to our old ideas we might have looked off into the distance and said, "There might be something there, but I would rather remain rooted to where I am." But because we decided to explore for a bit we've pushed ourselves out of the well and into multiple overlapping spaces of possibility.

1.1 THE PROBLEM OF PERSISTENCE

I have provided a number of links to videos of the movement of the creatures in this thesis. We still live in the early ages of hypertext, where we have not entirely solved the problem of persistence: how do I make a link to a document or file online that will remain accessible for more than a few years? This is a great concern for scholarly publishing, as stable references to material are necessary for continued development of a field.

I have decided to provide links to the materials of this thesis through a persistent uniform resource locator (PURL), a means of providing persistent links to networked resources. All information regarding the thesis should be available from `http://purl.oclc.org/NET/NKNOUF/MSThesis/`, with this document itself available from `http://purl.oclc.org/NET/NKNOUF/MSThesis/Thesis/MSThesis.pdf`. This represents an intermediate resolution service that I can update as necessary to point to new locations for the material. Importantly, however, the PURL will not change. This should hopefully at least partially solve the problem of persistence, assuming I keep the PURL updated throughout the years.

1.2 WHAT IS IN THIS THESIS

A document, like this thesis, is inherently linear. Even with the possibility of hypertext links between and amongst sections, the document is still structured in a fashion of chapter–section–subsection, repeat. One chapter follows another in an arbitrary organization, honed by some historical reasons perhaps known or not, and enforced by administrative and organizational rules that provide us with requirements. I write what is obvious in order to emphasize that I would prefer to do otherwise. Nothing about this thesis or my experiences with it has been linear. One idea in a book about science and technology studies lead me to rethink the design of my creature, which in turn lead me to think anew on how the creature fits into the discourse surrounding robotics. And in turn, I think about thinking about these things. Meta-thought, meta-ideas constantly interpenetrating with the first-order texts, objects, and programs.

I do not know what it would mean to create a text that was closer to these thought processes. I do not know whether it would be successful for the intense or casual reader. What I do know is that the construct of a book (or thesis) encourages us to think in a certain linear fashion—and that this is not even orthogonal, but entirely inconsistent with the actual production of this work.

Nevertheless, I provide here a type of *outline* of what is to come.

- *Voices Heard: Making Internal Experience Audible*. What *is* the voice, and how can we think about it? And what happens when we consider that which is not the voice, which falls outside the space of speech? Animals, speech, and the animal cry. Antonin Artaud and new types of vocal sounds. Radiophonic arts. New media installations. Recording the non-speech vocal sounds of others.

- *Influence of Objects on our Behavior*. How do we relate with objects in our world? And what happens when we start imbuing them with certain types of "intelligence"? Psychoanalytic approaches to the object. Computational objects and agency. Robotics—and alt-robotics.

- *Studying Novel Objects and Experiences*. What are new ways of conceiving technological development? And if we develop objects in these new paradigms, how might we study their use? Non-deterministic view of technology. Expanding our view of the possibilities of technology. Actor-network theory and how it applies to the thesis.

- *Puppetry and Marionettes*. What is the relationship between robotics and puppetry? And how can we draw from an old tradition for modern times? Puppetry and the performing object (and how it relates to our relationships with objects in general). The possibilities that are afforded by using puppetry. Robotic puppetry and marionette projects.

- *Design of the Robotically-Controlled Marionette*. What is the form of the creature, the robotic marionette? Where do we place it? And how does it move? Decisions made in the form of the creature. Earlier iterations of the design. The stage and lighting. Motor control. Physical simulation on the computer.
- *Design and Implementation of the Software Agent*. How do we decide between speech and non-speech? How do we control the robotic marionette? Signal analysis and machine learning. Transformation of signal features into movement through neural networks and interactive genetic evolution. Translation into movement in simulated and physical creatures.
- *Study of the Robotic Marionette in the Home*. How do people interact with the creature? A case study of the use of the robotic marionette. What we did not know before. Lessons learned.
- *Possibilities That Now Exist*. How does this thesis relate to other projects? What went wrong. What we would like to fix. What we can do in the future.

2

Voices Heard: Making Internal Experience Audible

…but words fail us when we are faced with the infinite shades of the voice, which infinitely exceed meaning. It is not that our vocabulary is scanty and its deficiency should be remedied: faced with the voice, words structurally fail.

Dolar (2006, p. 13)

What could be easier to understand than the voice? We all have one, we know its effect on others, we respond immediately to its absence. The voice has infinite gradations, yet we know how to name each one of those possibilities. The voice is intertwined in the most intimate way with the social construction of our lives; it acts as our external intermediary, making physical through sound our thoughts and our feelings.

Yet what could be more difficult to understand than the voice? Where does the voice end and unarticulated sound begin? Who possesses a voice? Do animal sounds contain the properties of the voice? And for those who cannot speak, do they have a voice? When we begin to push through the barrier of the word *voice*, a signifier that subsumes incredibly interesting and important particulars, we find that in actuality the voice is something we know little about. Sounds that come from our mouths are vastly more complicated than a single word can contain—this is most evident at the boundaries. While it is a futile task to police the boundaries, continually traversing them to discover what is on the other side enables us to discover much that might have been lost had we not looked over the fence.

In this chapter I want to consider some of these lesser-known edge conditions of the voice as they are the main consideration for this thesis. What happens when we live at the fringes of the sounds of the voice, when we enable others to push these boundaries themselves? We will see that to do such a thing opens up a space of expression that lies behind or beyond linguistic systems—and living in this space, even for a short period of time, enables emotional experiences that are more powerful than what occurs when we limit the voice only to words.

2.1 ANIMAL SOUNDS BEYOND LANGUAGE

"Bwarrrrrgh! Mwurrrrrgh! Grrrregh!"
"Waaaash! Shwurrrgh!"

Writing about the voice is like writing about music, an activity that is often doomed from the start. Because writing is intertwined with linguistics our marks on the page are limited to the phonemes that exist aurally. Cross-modally we are linked: the visual is attached to the aural. This becomes painfully obvious when we try to transcribe non-speech sounds. The grunts and cries that began this section were my feeble attempt to capture the sounds of the man-apes in the opening scenes of Stanley Kubrick's and Arthur C. Clarke's *2001: A Space Odyssey*. While each person who reads those lines can guess at the feelings that drove the production of those sounds, each person, in their head, will likely will hear the sounds differently. No matter how many combinations of consonants and vowels I use, I am still limited to consonants and vowels. And what happens when what we want to transmit lies outside that binary opposition? We miserably fail, the page becoming subservient to the voice and its shades that break through the sharp boundaries of the letter.

Moreso, the sounds of the man-apes shock us into our animal origins. Grunts, cries, screams, shouts, exclamations—these are our vocal heritage, not the staid assigning of symbolic thoughts to physical symbols. We must confront the sounds of the animal and question what we might have *lost* by confining much of our aural production to the sounds of language. Yet many works of philosophy create a sharp binary opposition between the sounds of animals and humans. Speech ⇔ non-speech, language ⇔ non-language. Always defined in terms of the *non-*, the absence of something. But what if it were *speech* that were the absence of something? If throwing away speech, at least for a moment, brought us to an area of existence that is forever cut off from us whenever we use language?

Many of the philosophical discussions of animals and their sounds focuses on policing the boundary between human and animals vis-à-vis language. In his masterful study of the place of animals in philosophy, literature, and film, Akira Mizuta Lippit shows the trend, starting with Aristotle, of placing speech at the top of a pedestal with animal "cries" far below. According to Aristotle's account, "not only does the expressive range of human speech exceed that of the animal's cry (which is limited to the two poles of affect, pleasure and pain), but speech establishes a large realm of communication" (Lippit 2000, p. 31). Descartes confined animals to mere automata, unable to participate in "genuine speech" (Lippit 2000, p. 33). Rousseau too kept the binary opposition between humans and animals on the basis of language, since "what the animal lacks is not intelligence but imagination, which is to say, language" (Lippit 2000, p. 39). For Rousseau language and imagination are intimately linked through the ability to know and sense death: knowledge of death requires the production of unseen images, and such an ability enables one to "perfect" oneself. Thus humans and animals are eternally split because of this lack of animals to produce language.

Lippit shows that not all philosophers are content to draw the line between humans and animals, however. According to Lippit, Derrida describes the animal cry as that moment that "pieces the world of language and the other" (Lippit 2000, p. 42), tracing this back to Burke and Hegel. In fact, he finds in Burke a turning of the animal cry on its head. Burke reasons that in the cries of the animal, "such sounds as imitate the natural inarticulate voices of men, or any animals in pain or danger, are capable of conveying *great ideas*" (Lippit 2000, pp. 43, emphasis in original). Lippit suggests that the sounds of the animal might directly reflect something of the emotion of which they represent, instead of being abstracted through the signs of language: "Without the semiosis that transforms sounds into words, animal utterances, like the *nonsense* of foreigners, can only portray the dynamic of affects and bodily states" (Lippit 2000, pp. 30, emphasis in original). We see this through the fictionalized account of the man-apes in *2001*. Even if we do not know the *exact* qualities of which the sounds represent, we at least understand the general shape of their experience. In this way language fails us; lack of knowledge of a language prevents us from having access to the emotional content of the utterance. It is the extraneous detritus of language, prosody, that is in fact beyond language, understanding without knowing the precise content of signs that we might not understand. In this way we can make a link between prosody and the cries of animals, as prosody reflects in some way the underlying emotional content of the utterance.

Lyotard brings us closest to the conceptual underpinnings of the thesis. According to Lippit's interpretation of Lyotard, from the "banished region" of animal cries "bursts of affect reenter the world through secret channels opened up by transference, or affective communication: transference opens a line of communication that is essentially antidiscursive" (Lippit 2000, p. 49). It is just these "bursts of affect" that I am attempting to enable through the design of *syngvab* . Creating a space where this "transference" can occur will hopefully bring people closer to a place of vocal sounds that are not just communication; not just something that is antidiscursive, but rather is a production of sounds without message, without the desire to understand a statement, but rather a place of feelings.

Before we move on to works that expand the possibility of the voice, I want to make a few notes on the qualities of the *human* scream. For, if "the elusive mythical scream was at the outset caused by a need, then it retroactively turns into a demand surpassing the need: it does not aim just at the satisfaction of a need, it is a call for attention, for a reaction, it is directed toward a point in the other whihc is beyond satisfaction of a need, it disentangles itself from the need, and ultimately desire is nothing but the surplus of demand over need" (Dolar 2006, p. 28). This translation of sound into need is most often found in someone in pain, for pain is unsharable, and it "ensures this unsharability through its [pain] resistance to language" (Scarry 1985, p. 4). Like we saw above with the man-apes of *2001*, and below with the work of radiophonic artists, language and its structure prevents the expression of large aspects of existence. In fact, the "physical problem of pain is simultaneously bound up with the problem of language creation", and there exist numerous questionaires for those in pain that *attempt* to provide the person in pain with a means of expression, through existing words, the magnitude of their experience (Scarry 1985, p. 7).

2.2 SPLITTING THE VOICE FROM THE BODY

The problematic of the animal sound is but the beginning of our questions of the voice. Even with the issues we just considered, the voice was still attached to an entity: the movement of the animal's mouth corresponded to the sounds we heard. With humans up until the beginning of the late-nineteenth/early-twentieth centuries this was also the case. Yet with the invention of radio and recording, and the ability to project sounds across time and space, this link between sound and living entity was broken. The absence of the link, as we will see, enables a whole host of options that otherwise did not exist.

It certainly must have been strange for the first listeners of the radio—a human voice coming from a cloth-covered grill, cut off from a physical body in the room. The warmth of a human was replaced by the warmth of electronic components, vacuum tubes glowing in the dark. But where was the face? The undulating lips concurrently outlining the shape of the sounds? The attachment of voice and body, kept intimately close for all of human evolutionary history, were torn asunder by the ability to project sounds across distance, cutting the tie between person and sound.

Allen Weiss, in his critical study of the radiophonic arts, describes how "radiophony transforms the very nature of the relation between signifier and signified [S / s]" (Weiss 1995, p. 7); this transformation is analogous to what we saw with animals, where the animal cry upset the sharp link between the utterance and the abstract symbol. As we will see in a number of cases, the ability to dismember the voice from the body via the medium of radio enables the production of sounds and experiences that exist outside of traditional linguistic structures. Language demands a fuzzy-precise connection between the sign and the signified, vague, perhaps, but a connection in any event. Radio and the experiments therein suggest a way to cut this tie and make an utterance float in the space in between and around sign and signified.

Weiss, in his discussion of the radiophonic artist Gregory Whitehead, describes Whitehead's "principia schizophonica" as a way to " set all of our voices out of tune, indeed out of body, in a psychoacoustic meldown or mix-up" (Weiss 1995, p. 80). Mirroring the schizo-analysis of Deleuze and Guattari, Whitehead proposes the creation of experiences, via the medium of radio, that split apart the traditional construction of a complete, whole, entire self. Cutting the voice from the body through its transmission over radio waves, crushing the linguistic nature of the voice through non-speech vocal sounds—Whitehead attempts to shock us into a re-examination of our basely quotidian view of the voice.

To do so, Whitehead created a radio piece sponsored by the Australian Broadcasting Company entitled *Pressures of the Unspeakable*, run by a fictitious "Institute for Screamscape Studies". Whitehead invited listeners to call a public number and have their screams recorded, the sounds of which were eventually broadcast. According to Weiss, through the *Pressures of the Unspeakable*, the recording and transmission of screams "reveals the chaotic depths of linguistic and vocal systems" (Weiss 1995, p. 83). "The scream may evoke the most profound phantasms as it shatters the coherence of the symbolic" (Weiss 1995, p. 84). Giving a voice to the scream, releasing the unheard sound, calls forth most powerful psychological feelings. The anonymity of radiophonic discourse, with the voice split from the body, allows such an experience to occur without the normal inhibitions imposed by society.

Besides Whitehead, we can also consider the relationship between the text of a piece of radio theatre and its eventual production. For as we saw in the beginning of the first section, there is an immense difficulty in transcribing the sounds of non-speech. Those who *do* make an attempt can be committed to the asylum, as in the case of Gaston Duf, whose "liberatory, potentially subversize" orthographic elements marked him as "insane" (Weiss 1995, p. 60). But these logorrhea, these orthographic gestures, are incredibly releasing, especially when actualized by the voice. The French playwright Valère Novarina plays with these ideas, imploring his actors to "reach and express this sub-liminal, corporeal core of speech" (Weiss 1995, p. 61). We will see shortly how this mirrors the glossolalic language of Antonin Artaud. Novarina attempts through his radiophonic work to create a "theater of the ears", "shattering the quotidian flow of language in a rare outburst of non-sense at all levels of discourse" (Weiss 1995, pp. 63, 66).

In a similar way, but outside the radiphonic arts, Kelly Dobson's *ScreamBody* allows the capture of a scream in a portable, somatic, sound-proof chamber (Dobson 2005). This "wearable body organ" can be kept close to the body throughout the day, only becoming used when necessary. A person can use the *ScreamBody* to capture a scream, later releasing it at a more opportune time or location. Must like Whitehead's radio work, Dobson's *ScreamBody* splits the voice, the scream, from the body in order to make the scream possible. Without the mediating effect of radio, or of the technological design of *ScreamBody*, the voice, paradoxically, would remain silenced, kept inside the body as a result of societal norm.

2.3 SOUNDS THAT MOVE BEYOND LANGUAGE

The fragmentation of the voice in these radiophonic works comes as a result of theoretical and practical thoughts from early in the twentieth century. As a result of the rise of industrialization there was much consideration as to one's place in the world, how one should situate oneself in relationship to the prevelance of machines. Some saw incredible emanipatory possibilities of creating radical new art, as in the futurist manifesto by Luigi Russolo, "The Art of Noises":

> Let us cross a great modern capital with our ears more alert than our eyes, and we will get enjoyment from distinguishing the eddying of water, air and gas in metal pipes, the grumbling of noises that breathe and pulse with indisputable animality, the palpitation of valves, the coming and going of pistons, the howl of mechanical saws, the jolting of a tram on its rails, the cracking of whips, the flapping of curtains and flags. We enjoy creating mental orchestrations of the crashing down of metal shop blinds, slamming doors, the hubbub and shuffling of crowds, the variety of din, from stations, railways, iron foundries, spinning wheels, printing works, electric power stations and underground railways (Russolo 2004 [1913], p. 12).

While the work of the futurists was influential, it is the writings and plays of Antonin Artaud that most directly impacted the path of this thesis. For Artaud wanted to completely rethink, *re-experience* the voice on stage, removing the reliance on the text from the theatre. No long would a play be based on what is on the page, what was *written*, but rather with the deepest psychological experiences of our lives, pull us closer to our core *as a living being*:

> To make metaphysics out of a spoken language is to make the language express what it does not ordinarily express: to make use of it in a new, exceptional, and unaccustomed fashion; to reveal its possibilities for producing physical shock; to divide and distribute it actively in space; to deal with intonations in an absolutely concrete manner, restoring their power to shatter as well as really to manifest something; to turn against language and its basely utilitarian, one could say alimentary, sources, against its trapped-beast origins; and finally, to consider language as the form of *Incantation* (Artaud 1958, p. 46).

In letters to friends and colleages about his *Theater of Cruelty*, Artaud questions the supremacy of language:

> It has not been definitively proved that the language of words is the best possible language. And it seems that on the stage, which is above all a space to fill and a place where something happens, the language of words may have to give way before a language of *signs* whose objective aspect is the one that has the most immediate impact on us (Artaud 1958, pp. 107, emphasis in original).

Thus, Artaud writes that "I make it my principle that words not mean everything and that by their nature and defining character, fixed once and for all, they arrest and paralyze thought instead of permitting it and fostering its development" (Artaud 1958, p. 110). This could be considered the defining point of the thesis: that in an exploration and encouragement of the non-speech we break out of what is "fixed" and enter into the flow of life that is compressed into the confines of words and symbols. Yes, language is, at the moment, necessary for communication with each other. But relegating us *entirely* to language and words blinds us to a whole space of possibile engagement with our lives. According to Deleuze this is made all the more worse by traditional psychoanalysis: "*All of psychoanalysis is designed to keep people from speaking and to take away the conditions of true expression*" (Deleuze 2006, pp. 84, emphasis in original). Artaud too was committed to the asylum, but this should not prevent us from creating a space, a place, where the explorations that Artaud suggests might be possible.

2.4 VOCAL MANIPULATIONS

Concomitant with the expansion of the voice as a realm of inquiry, both media artists and human-computer interaction (HCI) researchers have reconsidered the voice as not only a possibility for new modes of expression, but also as a tool for control of new interfaces. While speech interfaces have long been a dream of computer researchers and science fiction writers, non-standard vocal interfaces have recently come into vogue for use with mobile applications (Borden 2002).

We have seen, through our consideration of the voice so far, that non-speech vocal sounds offer gradations of sounds and experience that are not possible when we limit the sounds of the voice to linguistic possibilities. These smoothly varying changes can be used for the control of graphical user interfaces, as shown by Michael Murdoch in his thesis, where he created a system to move a mouse pointer on the screen through the use of "oooh" and "ahhh" sounds (Murdoch 2006). Related to this work is the Vocal Telekinesis project which provides a hybrid physical/digital

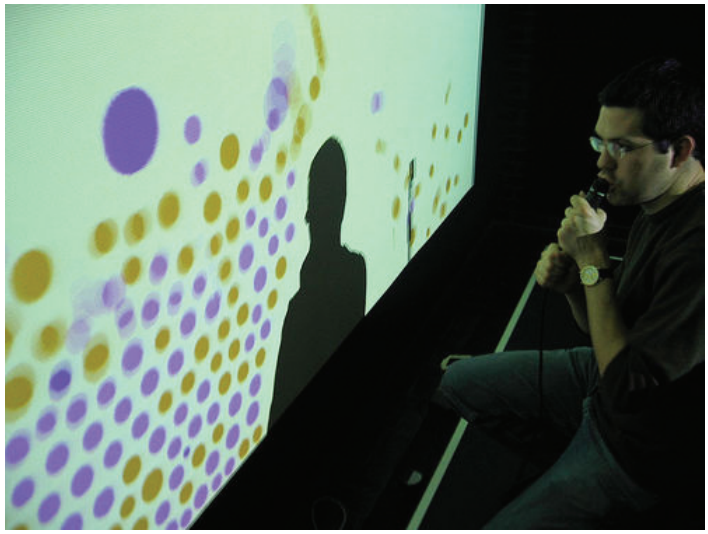

Figure 2.1: Installation view of Messa di Voce, by Golan Levin and Zachary Lieberman. Image from `http://flickr.com/photos/santheo/7392387/`.

interface and the control of inanimate objects through the use of the voice (Hashimi and Davies 2006). Additionally, vocal analysis of the type performed by these vocal interfaces can be used in learning applications for infants, as seen in the visiBabble project (Fell, Cress, MacAuslan, and Ferrier 2004).

Perhaps most relevant to this thesis, however, are the works of media artists. As part of the *Brain Opera*, a performance and installation piece by Tod Machover and the Hyperinstruments Group at the Media Lab, Will Oliver and colleagues created the "Singing Trees", a place for visitors to the installation to use their voice to try and modulate visual images by matching their singing pitch to a given cue (Oliver, Yu, and Metois 1997; Oliver 1997). Half a decade later, Golan Levin and Zachary Lieberman created two intertwined works that push the experience of the voice beyond merely the realm of singing. In the installation *Hidden Worlds*, Levin and Lieberman, through the use of augmented reality glasses, enable visitors to visualize the sounds of the voices by projecting mapped versions of the sounds onto a shared table (Levin and Lieberman 2004). In this installation, however, the visualizations can reference any vocal sounds, including speech (Figure 2.1). In a later work, however, they recruited two performers whose body of work involves extreme vocal manipulation. Their piece *Messa di Voce*, for the virtuosic voice performers Jaap Blonk and Joan La Barbara, was a live performance with real-time visual representations of their extended vocal abilities.

Drawing closest to our desires to "rethink" the possibilities of the voice is the work "BOOM" by Kelly Dobson (Dobson 2000). In this thesis project Dobson spent time with large construction machines, trying to negotiate her relationship with them, imitating their sounds with her voice. Dobson extended these ideas to the construction of a physical object, *Blendie* (Dobson 2004). I will speak more about *Blendie* in Chapter 3. Finally, Marc Böhlen's *Whistling Machine* was an attempt to encourage different types of expression in passersby, notably whistling, while focusing also on the ability of computers to "learn" a different type of "language" other than that of traditional semiotics (Böhlen and Rinker 2005).

2.5 WHEN/WHERE LISTENING BECOMES VOYEURISTIC

While we may know that many of the sounds that come out of our mouths fall outside the limited boundaries of language we rarely have occasion to listen to them with any intensity. We move through our days making all manner of bizzare hubhub, mostly oblivious to its character. These sounds are the detritus of the productive production of speech, the waste of our larynx slipping out of the corner of our mouths. Harumph's, meep's, ugh's...they go mostly unnoticed. So when and where might we confront these sounds? When they cause us embarrassment or concern—or when we create a situation that forces us to do so.

As I will detail in later chapters I designed *syngvab* to move only in response to non-speech vocal sounds as a way of enabling this sort of situation. While we have spent a large part of this chapter exploring different facets of the non-speech qualities of the voice, there is no *a priori* reason to assume that a computer will know what are the relevant structural components of non-speech vocal sounds. What we can do, however, is *teach* a computer program to make these discriminations, to, in a certain limited sense of the term, *learn* what aspects are different between the two classes. We do this through choosing from the family of machine learning algorithms, supervised and unsupervised means of generalizing from large examples of labeled classes of data. I will describe in much more detail in Chapter 7 the details of my implementation, but to foreshadow briefly: I am using a support-vector machine (SVM) classifier trained with a set of features calculated from classes of speech and non-speech examples I collected[1].

Through the collection of these sounds I discovered what was entirely unexpected when I began this thesis: that the production of non-speech sounds is an incredibly powerful and haunting experience. For matters of convenience at first, and later because of necessity, I began recording sounds from only close friends of mine. I would ask each of these people the same probing, direct, question: "Make sounds that express an emotional experience that you couldn't normally communicate through

[1] Of course, the SVM only knows speech and non-speech examples based on what I label, so any classification is entirely dependent on the decisions I make.

words." I would leave them in front of my laptop and exit the room, only returning when they were finished recording. This was done to given them as much privacy as needed and to hopefully make them feel as comfortable as possible with my strange request. Later, after I had left them, I could listen in the privacy of my office to the sounds they had made.

And as I listened, it was as if I were standing face-to-face with the person's psyche.

The voices were disembodied, detached from the person that produced them, much as we saw earlier in the examples of radiophonic transmission. They were also unlinked from any signifier, floating in a sea of possibilities absent any strong ties. I had to listen intensely to their emotional content, the waves of feeling that were imprinted on the sounds. In those moments of concentrated attention I felt deeps experiences of another person, projected my world onto what I heard, came closer to friends I already knew well.

When I had these same people listen to the sounds they had made they were amazed—not only at how personal the sounds were, but also at their ability to make those sounds. I too felt this; when I listened to the sounds I had recorded of my own it sounded, at times, as I were possessed. I was astonished at what I had been able to do.

These experiences were originally simply data-gathering expeditions, yet they profoundly illustrated a number of issues that are key to this thesis. First, the very *act* of making these sounds and listening to them is deeply personal and penetrating. When asked to do this, most people were hesitant, but in the end they found the experience rewarding in unexpected ways. Thus the process can have an impact that is entirely unexpected. Second, the view I had into these people's private lives underscored the need for a mediating presence in the project. I doubt these friends of mine, close as they are, would have been comfortable making these sounds with me in the room. It takes a certain acceptance on one's part in order to make sounds outside of language around others. This necessary distancing is one of the primary reasons why I am creating such a strange situation with *syngvab* . By having the robotic marionette move in response to these non-speech sounds I am hopefully lowering the barrier to the production of these sounds by removing a human presence that would otherwise be disabling. Paradoxically, however, I hope to show that such an interaction with a non-alive robotic creature brings people closer to some deep understanding of themselves.

⁂

Now that we have considered the possibilities of the voice, in what ways might we connect this to an object? We should first better understand the ways in which people form relationships with objects so that we can create a design method that uses the object to its fullest potential.

3

Influence of Objects on our Behavior

> The designers of computational objects have traditionally focused on how these objects might extend and/or perfect human cognitive powers. But computational objects do not simply do things *for* us, they do things *to* us as people, to our ways of being the world, to our ways of seeing ourselves and others.
>
> Turkle (2005, p. 1)

It is a platitude to speak of the importance of objects in our lives. We do not live in the philosophical dream world of purely mental states and thoughts—we are embodied in a physical environment. Not only are our thought processes, our neural activity, rooted in the firings of axons, our movement about the world, and our behavior in it, is linked to the objects around us. This chapter will focus intensely on the material world, and specifically the objects that inhabit it and the way we relate to those objects. How do objects affect us psychologically, how do they impact our mental processes? What happens to this when the objects, now termed *computational objects* begin to act on their own? When people start to form intense relationships with robotic objects, perhaps even preferring them to relationships with humans or animals? And what new objects might help to break free of the overtly representational vein that much robotics research finds itself in? By the end of the chapter we ought to be in a good place to see where *syngvab* fits into this space and to start asking questions about how to study interactions with the object.

3.1 OBJECTS AND OUR PSYCHOLOGICAL LIVES

The role of the object in psychoanalytic thought has often focused on the *human* as object, as opposed to the material, non-human, non-alive objects that populate our world. Now, choosing to begin with ideas from psychoanalysis might give some readers pause, especially if you are approaching this thesis from a scientific mindset. Thus, a few comments are in order before delving into the details of the ideas.

Psychoanalysis, especially the work of Sigmund Freud, Melanie Klein, and D. W. Winnicott, has an unfortunate reputation when viewed from a scientific perspective[1]. Science has proven these people incorrect, you might say. Indeed, the rise of cognitive and behavioral therapy, coupled with the development and prevalence of psychotropic drugs, has moved the traditional "talking cure" to the background. The results of cognitive science seem to show the brain as a set of mental "states", with the structure divided into a set of interlocking modules. Debates still exist as to the connections between the modules, how they might communicate with each other, and so on. However, the result appears to be unambiguous: the brain (not mind!) is relatively mechanistic and can be described in the context of modules, states, units and neurons.

We would be keen to be cautious, though, of any theory that purports to be "unambiguous". The brain, and our description of its activities as mind, may not be so easy to understand. Lest you think, from the previous sentence, that I subscribe to a dualistic theory of the workings of the brain, think not. The brain, as physical substance, and mind, as activities of physical substance, are two different ways of speaking of the same thing. And this is where we can make a connection to psychoanalytic thinking.

Since the brain and human behavior are such complicated and difficult to understand *things*, we might do well to consider them from a variety of perspectives. While cognitive science thinks through the mechanisms of how the brain works, psychoanalytic thinking helps us to conceive of the aspects of behavior that cannot[2] be quantified, as cognitive science demands. It is not that I see cognitive science as a "higher" or "better" description of humans and human behavior. Rather, it is a different *type* of understanding. "Higher" implies a ranking, a hierarchy, a division of description into discrete (or possibly overlapping) levels with definite direction. On the other hand, by framing the situation as different *types* of understandings, we can keep things separated that need to be separated. The various types of description can

[1] This is the case for many post-structuralist thinkers as well: much of Deleuze and Guattari's *Anit-Oedipus* is about how they see psychoanalytic thinking and its relationship to the repression or translation of desire as being dangerous to modern thought (Deleuze and Guattari 1983 [1972]).

[2] Perhaps at some point these things will be able to be translated into numbers, but what will the numbers mean? And will they tell us anything interesting about behavior?

remain incommensurate even while referring to the same thing[3]. They each tell us different (and interesting!) things about the brain and behavior. And they each can direct us in new ways of thinking.

FREUD AND THE OBJECTS OF THE MELANCHOLIC

To return to where we were. In psychoanalysis, *object relations theory* often refers to the *human* as object, rather than to how people relate to the material objects of the world, as you might expect by the name. Yet this can still help us to think about how we do relate to physical objects. In Freud's view people relate to others as objects most strongly in the experience of melancholia. In comparison to the relatively "healthy" experience of mourning after the loss of a person (either real, actual physical loss, or perceived loss), the melancholic turns the lost person into a perceived object; their status as human is, in turn, lost:

> We have elsewhere shown that identification is a preliminary stage of object-choice, that it is the first way—and one that is expressed in an ambivalent fashion–in which the ego picks out an object. The ego wants to incorporate this object into itself and, in accordance with the oral or cannibalistic phase of libidinal development in which it is, it wants to do so by devouring it (Freud 1953–1974 [1917] [1915], pp. 249–250).

For melancholics, the internalization of object (person) is a result of the redirection of mourning onto the subject himself. No longer does the melancholic look outward to the lack of the object in the external world; he turns inward, bringing the lost object inside and focusing on its outlines, for the object does not exist and is merely a void. All he can see is its boundary; its insides are empty.

Freud speaks not of the loss of a physical object (non-person), but we can speak of these events nevertheless. Think of the child and the teddy that gets left somewhere, or thrown away by accident. The object does not exist anymore, but the child still refers to it as if it is there. Teddy is internalized and its lack might overpower the memories of external interactions with it. Or think of Jack, the main character from the movie *Fight Club*. His apartment and all of his possessions are lost in a tragic gas explosion. The objects are gone, yet they still exist in negative form in his memory. This example as well points to the use of this idea in marketing and advertising. We see an object that we might want; we internalize the desire for it; we see the lack of it in our lives; we purchase the object to satisfy that lack; we see an object that we might want; and so on. The process draws directly from this mourning for non-existent or lost objects that Freud described.

[3] See the chapter "Multiple Worlds" in Law (2004, pp. 45–67) for much more on this topic.

WINNICOTT AND TRANSITIONAL OBJECTS

Returning to our childhoods for a moment, we can ask the basic question: when did I learn that there were things, *objects*, that existed independent of myself? There is no reason, *a priori*, for us to make this distinction, yet present introspection would show undoubtedly that the objects we interact with are external to our own bodies. So there must be some understanding to which a child comes early in her development by which she divides the world into exterior and interior reality. We need to dissociate this from the question of relating to other people such as our parents, and especially to our mothers; for while developmental and psychoanalytical evidence overwhelmingly suggests an intricate bond with our parents, where they exist in an in-between state of being both exterior and interior to us, there is a point where we discover that our parents are not us, that they are "not-me". But beyond this is our relationship with objects, a movement towards relating to objects as exterior to our inner reality, as "not-me" objects, things that exist independent of our parents and ourselves. To understand this would allow us to observe the development of inner and outer selves, and perhaps the formation of the self in particular.

D. W. Winnicott, in his essay "Transitional Objects and Transitional Phenomena", focuses on the events of the "first possession", calling objects "transitional" when they belong to an intermediate place, not entirely part of reality nor of fantasy:

> I have introduced the terms 'transitional objects' and 'transitional phenomena' for designation of the intermediate area of experience, between the thumb and the teddy bear, between the oral erotism and the true object-relationship, between primary creative activity and the projection of what has already been introjected, between primary unawareness of indebtedness and the acknowledgment of indebtedness ('Say: "ta" ').
>
> By the definition an infant's babbling and the way in which an older child goes over a repertory of songs and tunes while preparing for sleep come within the intermediate area as transitional phenomena, along with the use made of objects that are not part of the infant's body yet are not fully recognized as belonging to external reality (Winnicott 2005 [1971]*b*, pp. 2–3).

Transitional objects are items that "are not fully recognized as belonging to external reality": they present a difficulty to the binary division of the world into exterior and interior. Transitional objects, in the case of the infant, are an early means of dealing with the world, of relating to external reality.

Transitional objects are not meant to last into adulthood. Indeed, the fate of the transitional object is for it

> to be gradually allowed to be decathected, so that in the course of years it becomes not so much forgotten as relegated to limbo. By this I mean that in health the transitional object does not 'go inside' nor does the feeling about it necessarily undergo repression. It is not forgotten and it is not mourned. It loses meaning, and this is because the transitional phenomena have become diffused, have become spread out over the whole intermediate territory between 'inner psychic reality' and 'the external world as perceived by two persons in common', that is to say, over the whole cultural field (Winnicott 2005 [1971]*b*, p. 7).

Transitional objects are simply what their name suggests: a transition in the course of development from an understanding of the world where there is no difference between internal and external reality, to one where the infant begins to understand this difference. In "The Use of an Object", Winnicott says that, "At the point of development that is under survey the subject is creating the object in the sense of finding externality itself, and it has to be added that this experience depends on the object's capacity to survive" (Winnicott 1969, p. 714). In fact, to keep hold of the transitional object into adulthood, for it to not "lose its meaning", is a sign of psychosis: for if the adult is continually "finding externality itself", she has not learned how to integrate her internal reality with what she experiences externally.

Winnicott relates the experience of transitional objects to the "potential space": "The place where cultural experience is located is in the *potential space* between the individual and the environment (originally the object). The same can be said of playing. Cultural experience begins with creative living first manifested in play" (Winnicott 2005 [1971]*a*, p. 135). Thus the adult looks for the transitional object during her experiences with culture, attempting to return to the space where her relationship with reality was ambiguous, where there was still potential.[4] So relationships with cultural activities as an adult can have a transitional-phenomena-like quality to them, all the while attempting to bring people back to the first possession.

Transitional objects, as applied to the healthy individual in adulthood, can bring us to these "potential spaces". Development of new objects that place us in these intermediate spaces, where I do not know where I end and the object begins (and I *know* that this is occurring), pushes us beyond the purely functional interaction we have with many objects in our lives. The object is less a tool, subservient to my control and desires, but plays more of a role in the interaction. If there is a way to design for this, then we should do it! I do not make the strong claim that the object I developed in this thesis actually *is* a transitional object. I suggest that it might be, however; that it might perturb us from the purely binary division between subject

[4] "This potential space is at the interplay between there being nothing but me and there being objects and phenomena outside omnipotent control" (Winnicott 2005 [1971]*a*, p. 135).

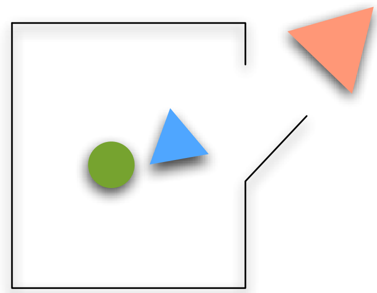

Figure 3.1: Recreation of an image from the motion and agency studies of Heider and Simmel. The large triangle is often described as "bullying" the smaller triangle and its "companion", the circle. Based on Figure 1 from Heider and Simmel (1944).

and object, and bring us, as described in the Introduction, away from the poles of dualisms and into the darkened spaces around them. Indeed, transitional objects in adulthood are all about opening up new spaces for experience and possibility.

3.2 AGENCY AND COMPUTATIONAL OBJECTS

Children have oft had a fungible boundary between what is alive and what is not; what can be considered to have *agency* and what is merely another object in the world. We learn as adults to make this boundary concrete; to ensure that what we *know* to be alive (given the biologic definition) remains separated from merely "objects" that *cannot* be alive (because they do not posses goals, desires, wants, feelings, and so on). This division comes as a result of development in the child. The relationships with the objects of the environment, either a real or constructed one, have been well-studied paths. Indeed, Piaget has shown how the simple movement of an object can be enough to get a child to view it as "alive", noting that depending on the level of development, even a rock can be "alive" if it is seen to move (Piaget 1951 [1929]). Adults too can be made to attribute agency to the inanimate, even to the not-embodied. The early work of Heider and Simmel showed how observes ascribe intentionality, desires, and goals to geometric objects moving on a screen (Figure 3.1) if the path of these objects follows an expected trajectory (Heider and Simmel 1944). This result has been shown in a variety of different experiments over the past fifty years or so, with much work being done to consider the relationship between these simple visual displays and the complicated and higher-level social and causal processing that is also going on (Scholl and Tremoulet 2000).

But the construction of computational objects upsets this distinction, not only

Figure 3.2: Paro the robotic seal, designed for use by the elderly in nursing homes. Image from `http://flickr.com/photos/granick/205552836/`.

for children, as you could imagine, but for adults as well[5]. And it is this perturbation of the human/object system that is most interesting for this thesis and my work. If we want to possibly create objects that push people into the "potential spaces" of Winnicott, what better place to begin than with objects that upset the distinction between you and them? Maybe there is a way in which this boundary-hopping enables movement into the expanded space of cultural experience suggested by transitional phenomena. Perhaps new types of objects, like computational objects, are better suited for boundary-blurring.

This certainly seems to be the case, at least according to many observations of researchers over the past decades. Sherry Turkle has extensively studied the responses of children, and adults, to new technological developments, including the proliferation of computational objects. Her work has directly shown how objects that seems to act own their own cause both children and adults to question basic ideas of aliveness, agency, and intention. For children, the "issues raised by a smart machine spoke directly to philosophical questions already on children's minds" (Turkle 1995, p. 82), that is, what it means to be alive. Adults, however, were more able to keep the boundary intact, falling back on the statement that it was "just a machine".

[5] The object does not even have to be representational for the distinction to blur. Nass and colleagues, among others, have detailed the many ways in which people will respond to their computers as if they were alive in a "mindless" manner (Nass and Moon 2000). He calls this attribution of agency and intent *ethopoeia*, meaning referring to an object as human while knowing that it does not warrant this action.

Interactions with these new types of objects is not limited to children and adults, however. Now the elderly too have to negotiate the concerns raised by robotic creatures. Much current robotics research focuses on the use of robotic creatures as companion animals in the nursing home as a result of the increase in the number of people aging in Western societies. One such robot is Paro, the robotic seal (Figure 3.2). Turkle and colleagues have extensively studied the interactions of the elderly with Paro (Taggart, Turkle, and Kidd 2005; Kidd, Taggart, and Turkle 2006), suggesting the possible danger of reliance on the robot as a stand-in for interactions with people. Indeed, there is a certain "seduction" to the use of Paro to solve the problem of lack of caregivers (Taggart, Turkle, and Kidd 2005, p. 61), yet this "solution" merely raises the concomitant question: why are we spending research money and time on a robotic companion instead of spending the same money on training human caregivers instead?

This represents the concern of Turkle and others about the deep, personal relationships that we form with these objects, these "relational artifacts". Whether it is evident to the robotics designer or not, relational artifacts are able to push our "Darwinian" buttons, especially for children or the elderly (Turkle 2006). They bypass our rational sensibilities, speaking directly to our concern for others, our desire for love. Turkle warns that

> Relationships with computational creatures may be deeply compelling, perhaps educational, but they do not put us in touch with the complexity, contradiction, and limitations of the human life cycle. They do not teach us what we need to know about empathy, ambivalence, and life lived in shades of gray. To say all of this about our love of our robots does not diminish their interest or importance. It only puts them in their place (Turkle 2006).

Nigel Thrift, in his review and questioning of "electric animals" also worries about the perturbation of subject and object that comes with blind application of biological metaphors in new robotic creatures, raising the the oft-neglected *ethical* concern regarding our relationships with the not-alive:

> …the advent of software-driven entities modelled on biological assumptions is a significant event that has the potential to decisively change everyday life by adding in a new range of cohabitees. In particular, it offers a new set of ethical dilemmas that have clearly not been solved in the case of companion animals (Thrift 2004, p. 476).

Perhaps, however, if we were to move away from the representational we could allay some of these concerns and use these "Darwinian Buttons" to investigate other

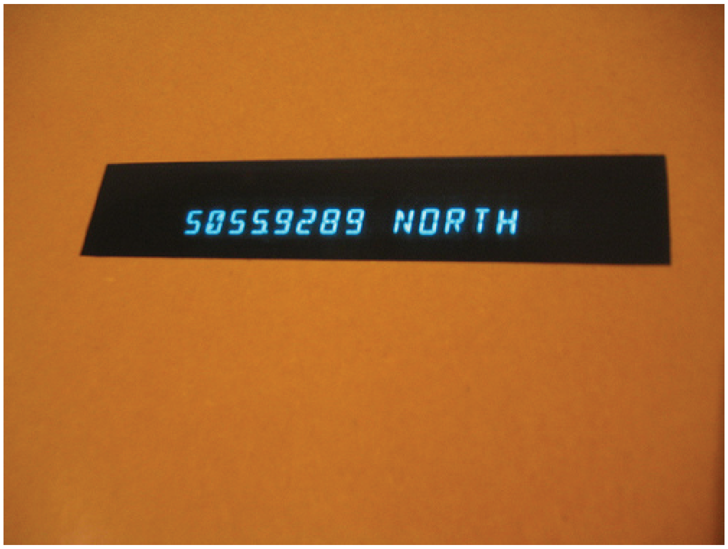

Figure 3.3: Installation of the GPS Table by Dunne and Raby, from their Placebo Project (Dunne and Raby 2002). Image from `http://flickr.com/photos/nearnearfuture/425450181`.

questions of agency and relationship that are not so immediately frightening. The artists Anthony Dunne and Fiona Raby, in their aptly-named *Placebo Project*, created a series of objects to address people's concerns about electromagnetic radiation. None of these objects resembled animals or people in any sense, yet the responses by participants reflected at least some conception of the object as possessing agency or desires. In the *GPS Table* (Figure 3.3), a screen embedded in the tabletop displayed the GPS coordinates of the object, or the word "lost" if it was unable to make a connection to the GPS satellites. The response by the family to this simple programming was fascinating. The father reacted with surprise to his "concern" for the well-being of the table:

> "When you came and installed it and we plugged it in and it said 'lost', I was absolutely shocked. I'm not quite sure why I was shocked. I thought 'Bloody hell, the poor thing's lost.' It's a clever use of the word, because clearly it starts emotions with us. It's a very emotive word, and it's interesting to work out whether you could extend the *vocabulary of the table*, because that would be great fun. Things like 'trying to find home'. You don't have to have a huge vocabulary. It could be just five words" (Dunne and Raby 2002, emphasis added).

Important here to also note is the father's desire create a "vocabulary of the table", suggesting the ways in which objects, especially computational ones, *speak* to us. The father might be indeed referring to actual speech or orthographic marking, but I see this as reflecting the deeper power of an object to affect us in entirely unexpected ways, "talking" to us in a non-linguistic manner, communicating something powerful.

We can consider the example of the GPS table as an *uncanny* situation. For Freud the uncanny is "that species of the frightening that goes back to what was once well known and had long been familiar". (Freud 2003 [1919], p. 124) But the uncanny is not always frightening: it can also represent a resemblance to something we know but can not entirely remember. Especially when we create new types of robots is the uncanny useful, as it "deals explicitly with the boundaries between the natural and the artificial, the animate and the inanimate, the human and the nonhuman" (DiSalvo 2004, p. 4). Drawing from the work of Freud and this concept we can directly connect with a body of theoretical work in the humanities that have used the uncanny as a way to understand a variety of theoretical concerns.

This question of interaction with objects, of how we ascribe agency, desire, and intent to the inanimate, and how we should interact with them (an ethics), represents a broader concern about the role of objects in societies. Especially in the case of experts and their peculiar interactions with the objects of their work, Karin Knorr Cetina has observed that "the libidinal, reciprocal and in other ways binding components of experts' object ties make it plausible to construe these relationships as forms of sociality rather than simply as 'work' or 'instrumental action' " (Cetina 1997, p. 23). Objects are not merely tools to instrumental ends, as she has observed; rather, for the expert, and, I would argue, for many of the users of computational objects as well, there is a "sense of bondedness or unity (an identity feeling) with objects, a moral sense (the oughtness of approaching them in certain ways), and states of excitement reaffirming the bondedness" (Cetina 1997, p. 20). These bonds with objects, while important and necessary for the expert, can become difficult for the rest of us. Fractionation is a result, a trait of the posthuman as described by N. Katherine Hayles. Hayles suggests an emancipatory power to this splitting of ourselves in and amongst the world:

> We are never only conscious subjects, for distributed cognition takes place throughout the body as well as without; we are never only texts, for we exist as embodied entities in physical contexts too complex to be reduced to semiotic codes; and we never act with complete agency, just as we are never completely without agency (Hayles 2001, p. 158).

The development of computational objects thrusts us into the posthuman world, removing boundaries that have existed with some strength for many years, and forcing us to come to terms with our distributed selves.

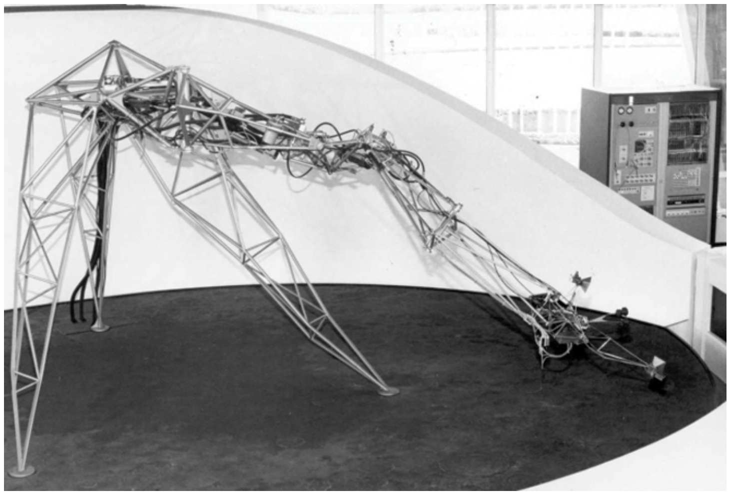

Figure 3.4: The senster robotic creature by Edward Ihnatowicz, from the early 1970s. Image from `http://www.dse.nl/~evoluon/senster-e.htm`.

3.3 ALTERNATIVE WAYS OF THINKING ABOUT ROBOTIC CREATURES

Robotic creatures do not need to have a representational form to provoke strong reactions in people. The reliance on *what is already there* causes the developers of robotics to merely refer to existing creatures and standard biological metaphors, forgetting that when we create our own beings and objects we do not have to remain rooted to reality. Designers do not have to refer to existing animals or humans as a crux to break down the initial barrier between person and novel object. It is in fact possible for people to have immediate reactions to new creatures, new objects, new experiences, without there being a direct representational connection. Yes, people will form their own links between what they see and their prior experience, but this does not have to be made explicit by the robot designer[6].

One of the earliest projects to draw from this stance was *Senster* by Edward Ihnatowicz (Figure 3.4). The movement of the creature was extremely simple by today's standards: it would move closer to soft sounds, away from loud sounds, and away from your (attempted) touch. But the responses by visitors to the installation were extremely complicated, belied by the basic control mechanism. The unpredictability of the acoustic environment and public coupled to produce sophisticated, "life-like" movements of *Senster* . The robot could be said to have an "arm" and a "body", but

[6]This is the case even when the goal is not necessarily *interaction* with the robot, as in the case of service robotics (Forlizzi and DiSalvo 2006).

Figure 3.5: Petit-Mal by Simon Penny. Image from http://ace.uci.edu/index.php/ACE/proj/12/.

they were not directly representative of any existing creature, real or imagined; yet the responses by the public suggested that this did not matter, that they could interact with the object as if it were a creature.

In common with *Senster* , the control system for robotic creatures does not have to be "intelligent" in any way for people to respond to the creature in an intense, personal manner. In fact, the robot can be as simple as an inverted pendulum with a basic feedback control system, as in *Petit-Mal* (Figure 3.5) by Simon Penny (Penny 1997; Penny 1998). His piece, when installed in museums, caused people to spend many hours trying to understand the "motivations" and "desires" of the object—even though there were none "programmed" into the system: "Another curious quality of Petit Mal is that it trains the user, due to their desire of the user to interact, to play; no tutorial, no user manual is necessary" (Penny 1998, p. 464). No coding of states, nor "understanding" of the world; rather, the blind application of a control system designed to keep the two parts of the pendulum in equilibrium. And yet, people responded with such fervor as to believe, from their point of view, that there was something "intelligent" controlling the object. Says Penny: "The machine is *ascribed* complexities which it does not possess. This observation emphasises the culturally situated nature of the interaction" (Penny 1997, p. 105)[7].

Kelly Dobson's humorous objects unsettle our complacent attitudes towards objects of our daily life. Most important relevant for this thesis is her project *Blendie* , a common household blender that only turns on when you make the sound of the object (Dobson 2004; Dobson, Whitman, and Ellis 2005). The project could perhaps be seen as unnecessary: why would we want to interact with an object in this manner? Yet it directly confronts the commonly held desire that interaction with objects should be done in a purely functional manner. Why should it not be amusing, fun? Why should the object be subservient to us? Why should we not "speak the language of the blender"? Dobson's work forces us to think of these questions, to consider the reasons behind their asking.

Her current work is also in this vein. In the *Wearable Body Organs* series, she considers creating objects that help address the unmet needs of people—those that are unvoiced and unknown (Dobson 2005). For example, in *Scream Body* , Dobson has created an object that allows us to release pent-up screams whenever and wherever needed. Her most recent project, *Omo* , is an egg-shaped object that breathes with you—that reflects your internal state, perhaps helping to make you relaxed, or perturbing your breathing (and consequently your physical state) when desired. The object is held close to your body, supporting the deep somato-sensory connection between breathing and ourselves. Again, the work of Dobson suggests an entirely different way of relating to objects, and their creation, than the purely functional manner commonly envisioned by traditional HCI.

Recently, the media artist Marc Böhlen has been exploring objects and agents that do not necessarily behave, that have a "life" of their own, responding in complicated ways to people and their environment. His robotics work explores the concept of *contemplative robotics*, a more complicated way of thinking the human-robot relationship:

> The richness in current technological innovation should be coupled explicitly with awareness of the cultural contexts in which it operates. We can do more than build efficient machines. We can create robots that are acceptive of human folies and fears, robotic systems that understand that the human notion of time and duration is different from clocked machinic time. Expanding the robotics agenda in such ways can help develop better robots to live with in the long term (Böhlen and Mateas 2002).

[7]Some traditional robotics researchers are working to understand this situatedness and pull from it in their designers and development; see especially Breazeal, Buchsbaum, Gray, Gatenby, and Blumberg (2005), Lee, Lau, Kiesler, and Chiu (2005), (Michalowski, Sabanovic, and Michel 2006), and (Michalowski, Sabanovic, and Kozima 2007).

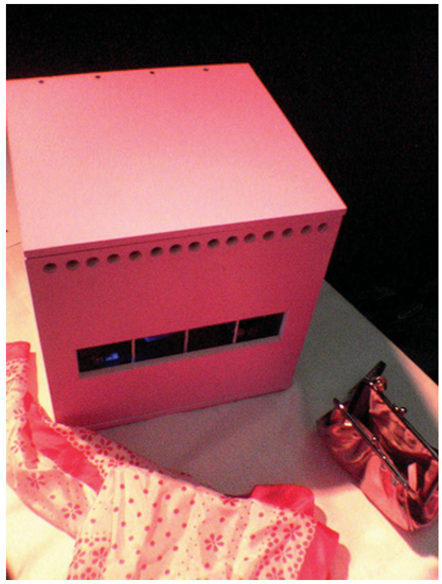

Figure 3.6: Image from the Amy and Klara project of Marc Böhlen. Image from http://flickr.com/photos/youraccount/225039093/.

His recent project, the two robotic agents Amy and Klara (Figure 3.6), throw "hissy-fits" and frankly do not like each other (Böhlen 2006). Their response to the environment (the other agent) is imperfect, due both to the noisiness of the surrounding area and the difficulty of speech recognition. The project is an attempt to subvert the belief that agents, especially ones that are embodied, must be subservient and responsive to our every command. What would it mean to create an agent that talked-back to you? Perhaps this is desirable, as a human that always follows our beck and call is not enjoyable to be around. Böhlen's project is a way of thinking about a more complicated way of interacting with the objects of our creation, given the possibilities we now have for creating these things.

* * *

What we should take from this chapter is that the interactions between people and objects is much more complicated than many in the field of HCI and robotics might like to admit. We cannot assume that people are unified wholes; that they act towards objects in a rational manner; and that perceptions of agency require artificial intelligence algorithms. This is not to say that we should assume that people merely revert to a child-like state when these assumptions are validated; people *know* that an object is not alive, they know about how the world works, yet there is a strange juxtaposition that disturbs the situation. Perhaps this is merely a transitory time, that at some point in the near future we will know better how to deal with these new classes of objects, better predict when and how we will respond to any particular object with such an intensely personal reaction. But perhaps not. Maybe this is our new life with objects—difficult to pin down, impossible to predict, always interesting and powerful enough to break us out of routine. The new objects demand a new way of study, one that does not try to fit them into pre-existing categories and respects the quantitatively and qualitatively new qualities.

4

Studying Novel Objects and Experiences

> Method? …It is also, most fundamentally, about a way of being. It is about what kinds of social science we want to practise. An then, and as a part of this, it is about the kinds of people that we want to be, and about how we should live.
>
> Law (2004, p. 10)

If there really are new objects in midst, if we assume that these objects are perhaps both *qualitatively* and *quantitatively* different from prior types of objects, and if we believe that they influence our behavior in different ways, in what way should we study our interactions with them that illuminate their novel aspects? Maybe a traditional user study will conceal what is most interesting, rather than reveal what might be new. What new ways of studying science and technology might we be able to draw from in order to study these interactions? This chapter will talk about some of the problems of traditional understandings of science and technology, and quickly move into different ways of conceptualizing the practice of technology development. This transitions into different ways of *studying* science and technology. I want to suggest that these new ways of studying things work in our favor and might be a better way of figuring out how the new types of objects we explored in the previous chapter influence our behavior. I want to suggest that the methods I describe in this chapter might be just the beginning of a new methodological enterprise.

4.1 QUESTIONING THE ASSUMPTIONS OF TECHNOLOGY

Technology is not neutral.

Perhaps you see this statement as obvious, as merely reflecting the current understanding of the relationship between technology, society, and politics. Or perhaps you read this as a direct challenge to your beliefs about technological development, your understanding of the split between, on the one hand, the results of technology, and on the other hand, the personal and social qualities that then control technology. However, my stance in this thesis is to ally myself with the former view. This section will review the arguments in its support, and my hope is that if you subscribe to the latter you will at least acknowledge the merits of the former.

Since perhaps the early 1970s there has been a trend in the humanities and social sciences for direct study of technological development and its relationship and dependence upon personal and social factors. The main result of this body of work, if we can choose one single point, is that technology and society are intimately linked: you can't study one without studying the other. Technology does not exist independent of a society, and, especially in modern Western ones, societies function as a direct result of technological development. This symmetry is opposed to the commonly-held belief that technology and its development is in some way prior to and independent of its home society.

There are many texts that address these issues from the non-deterministic perspective. I will choose to focus on one, however, that offers a particularly lucid account and approaches the problem from a distinctly critical perspective. Andrew Feenberg, in his book *Questioning Technology*, provides a convincing set of arguments for rejecting many of the commonly-held beliefs about technological development, such as determinism and substantivism. Technological determinism is the belief I rejected in the opening of this section: that technology is merely a neutral means for productive and useful ends. Or, to be phrased in the negative sense, technology exists for the purposes of domination and its use will eventually doom us[1]; this view is also termed *substantivism*, suggesting the existence of some negatively-tinged "essence" of technology that we are at a loss to excise (Feenberg 1999, pp. 1–3). These dystopian views of technology, while indeed not remaining subservient to the "technology is neutral" ethos, instead push us dangerously close to the Luddite and reactionary, one who wants to stop technological development at all cost. Rather, Feenberg suggests that we can first accept the ideological implications of technology. To do this "has definite political implications. If one can loosen up the public vision of technology, introduce a contingency into it, technical elites will have to be more responsive to a democratically informed public will" (Feenberg 1999, p. 8). By taking a "social account of the essence of technology [that thus] enlarges democratic concerns", Feenberg thus rejects the "gloomy Heideggerian prediction of techno-cultural disaster" (Feenberg 1999, p. 17).

Even so, it would behoove us to observe, without the need for much reflection,

[1] This is often the view when speaking of the horrors of nuclear warfare.

that we live in what can be called a technocracy, the "generalization to society as a whole of the type of 'neutral' instrumental rationality supposed to characterize the technical sphere" (Feenberg 1999, p. 75). Technocracy is not compatible with democratic institutions for obvious reasons: we remove the guidance of the electorate, or citizenry, and replace it with the guidance of a few non-elected technological leaders. They ensure us that we are progressing smoothly, that the tools they are developing and releasing for our use are decidedly "neutral" and for our benefit. These leaders will tend to select technologies that support existing hierarchies and that retain the hegemonic systems of control, overt or covert (Feenberg 1999, pp. 76,86–87). Of course, this does not have to be the case, and the bit that I have outlined is perhaps too simplistic. Indeed, as we will see below, there are a number of researchers whose technical practices tend to subvert existing ideas of traditional technological development. Feenberg calls this *reflexive design*, writing that a "reflexive design process could take into account the social dimensions of technology at the start instead of waiting to be enlightened by public turmoil or sociological research" (Feenberg 1999, p. 90).

This challenge to the technocracy suggests that we have to consider the problem of agency[2]; that is, what possibilities are there for those who are not members of the technocracy to influence technological development. Here we can draw from numerous critiques of large social systems, or as Feenberg terms them, "networks". "Systems, as self-reproducing wholes, are fragile subsets of much more loosely organized complexes of interacting elements that may support several overlapping systemic practices" (Feenberg 1999, p. 118). This suggests the potential power of techniques that subvert these networks, actions that focus on weak or ignored links in the network, in line with the "tactics" of Michel De Certeau (Feenberg 1999, p. 112). Additionally, the use of participatory design, a technique also used by reflective HCI researchers (see below), enables agency to be distributed beyond the confines of the technocracy and given to the democratic citizenry (Feenberg 1999, p. 123).

4.2 CRITICAL TECHNICAL PRACTICE AND REFLECTIVE HCI

If this is the view from the outside, of a sociologist or philosopher trying to understand techno-social development through the process of observation, what mongrel might be unleashed if we add practice to the mix? If there were to be some method of approaching the development of new technology with the knowledge of hermeneutics, social science, and philosophy? Recently this has been happening through a variety

[2]I want to note that this use of the term agency is subtley different than my earlier uses. I am not speaking here of the human's perception of whether or not an object acts own its own or possesses some form of agency; I am speaking rather of the traditional use of the term, of the ability to enact change through the actions of a human actor.

of methodologies, most notably that of *critical technical practice*. Critical technical practice as a term was first coined by Phil Agre as a response to his experiences as an AI researcher in the late 1980s and early 1990s (Agre 1997). He observed the development of AI technologies with a dual personality: one part, AI researcher and creator of new technical methods; the other as a reader of social science and philosophy with the goal of understanding their application and, indeed, *usefulness*, for technical work. Agre used the techniques of textual interpretation drawn from the humanities, known as hermeneutics, and applied them to the discourse of AI to show how the words used by AI researchers helped determine the means of solving problems and the outcomes of disputes: "AI's elastic use of language ensures that nothing will seem genuinely new, even if it actually is, while AI's intricate and largely unconscious cultural system ensures that all innovations, no matter how radical the intentions that motivated them, will turn out to be enmeshed with traditional assumptions and practices" (Agre 1997, p. 151).

This point about "reading" the discourse of AI might seem entirely foreign to those approaching the thesis from a purely scientific or technical bent. But the point, while subtle, should not be passed over. Agre, among others, shows that the words used by researchers, and the way they use those words to enable to deny the presentation of ideas, counter-ideas, solutions or refutations, directly impact the types of projects or ideas that the field, as a codified unit, will deem important. With the idea of critical technical practice, Agre suggests a different route, one that uses the slipperyness of language, its inability to perfectly, mathematically capture all of the nuances of an idea, to instead reflect in a critical manner on the work of the field. Agre notes the difficulty of this endeavor, but suggests that the outcomes will be entirely worth the effort:

> A critical technical practice will, at least for the foreseeable future, require a split identity—one foot planted in the craft work of design and the other foot planted in the reflexive work of critique. Successfully spanning these borderlands, bridging the disparate sites of practice that computer work brings uncomfortably together, will require a historical understanding of the institutions and methods of the field, and it will draw on this understanding as a resource in choosing problems, evaluating solutions, diagnosing difficulties, and motivating alternative proposals. More concretely, it will require a praxis of daily work: forms of language, career strategies, and social networks that support the exploration of alternative work practices that will inevitably seem strange to insiders and outsiders alike. This strangeness will not always be comfortable, but it will be productive nonetheless, both in the esoteric terms of the technical field itself and in the esoteric terms by which we

ultimately evaluate a technical field's contribution to society (Agre 1997, p. 155).

Since the publication of Agre's work many researchers in fields as varied as AI, psychology, human-computer interaction, robotics, sociology, and art have begun to develop their own critical technical practices. There is not room here to detail the work of all individuals or groups, so I will focus mostly on that of Phoebe Sengers, Bill Gaver, and colleagues. Their research has taken the ideas of Agre and applied it directly to the development of new technological objects that spring forth from a critical engagement with technology and thought. In an early project, Sengers created an alternative to the prevailing view in HCI, then and now, that emotion and feeling should be structured and codified, divided into mathematical states. Called the Influencing Machine, the project was a way to focus on the "enigmatics" of affect, creating subjective experiences that enable interpretation and resist easy classification into arbitrary categories (Sengers, Liesendahi, Magar, Seibert, Müller, Joachims, Geng, Maartensson, and Höök 2002). Later work by her group has extended these methodological and conceptual ideas to a number of projects, suggesting direct applications to HCI design (Boehner, David, Kaye, and Sengers 2005). This leads to the concept of *reflective HCI*—reflective not just for the designer, as he makes decisions in the development of a new object or interaction, but also for the user, creating a space for her to be free to be *human*, refusing to be codified into a discrete set of states selected by a distant designer. Reflection using technology and the reflection of alternatives, the possibility of other options not envisioned by the designer, but enabled by not shutting off prospects unnecessarily (Sengers and Gaver 2005; Sengers, Boehner, David, and Kaye 2005; Sengers and Gaver 2006).

Other researchers in the HCI field have also explored the possibilities of open interpretation, and the consequent ability for abstraction, in the design of technical objects. These projects all draw from an understanding of the need to consider the social aspects of object use—not just in an off-hand way, but to really get oneself *dirty* in the details of use by placing the objects within an existing space, removing the abstracting nature of the laboratory for the messy real world (Mutlu 2006; Mutlu, Forlizzi, Nourbakhsh, and Hodgins 2006). For a consideration of the "social" will remain only theoretical as long as you define the social to be what can be found in a space you control. The use of the objects we design must be observed in as *in-situ* of a situation as possible—spaces where we cannot predict values for all of the variables, or places where "variables" are a foreign concept.

4.3 ACTOR-NETWORK THEORY

We've seen how traditional assumptions of technological development and power need to be rethought, that we cannot continue with the same concepts that assume a neutral means for the given ends. And we've seen that when we begin to understand our work in this way, we come up with different metrics and methods for the creation of technological objects, focusing more on the expansion of possibility rather than *a prior* limitation to a small subset of options.

But if we continue on this path of pushing away from the assumptions we have no rejected, we come up with our final hurdle: that of the study of the object. Researchers in reflective HCI and related fields already use *ethnomethodology*, a technique of close participant and user observation. This already is anathema to those versed in the cognitive-psychology-influenced methods of the traditional user study. But perhaps we need to rethink even this; perhaps the entire enterprise of observation and studying people needs to be rethought and re-conceptualized in order to take into account similar texts and ideas that have influenced our ideas so far.

This is the work of the *sociologists of scientific knowledge*, sociologists of science who try to understand the means of scientific knowledge gathering. Through a number of detailed *in-situ* studies of the *practice* of knowledge creation, we can see how the creation of new technical ideas is dependent on a host of objects, people, and methods that have hitherto been ignored.

"If this is an awful mess ...then would something less messy make a mess of describing it?" (Law 2004, p. 1) John Law, a sociologist of scientific knowledge, is describing a juxtaposed drawing: images, geometric shapes, text, lines. No "order", as we would usually understand it. His use of the graphic illustrates the messy situations of contemporary sociological studies of technology. When you consider the variety of actors in the creation of any one technological artifact, the incongruence of each with the other becomes apparent: graduate student, advisor, data sheets, assembly code, institution, users, local community, conference paper, *etc*. Would a framework that attempts to smooth away the differences between each of these actors merely "make a mess" of the situation? Would purification into the oft-mentioned spheres of nature and society (Latour 1993) really give us a better understanding of these novel experiences? No; indeed, as Latour so well describes, this split into seemingly independent areas of "nature" and "society" is itself arbitrary, a historical development the demise of which we are finally beginning to witness. When we begin to create objects that already work to blur the boundary between person and object, we are immediately faced with situations that bring these spheres (un-)comfortably close together, leading to the creation of more and more "quasi-objects" (Latour 1993).

Latour, Law, and others have developed a way of thinking about the study of these contemporary situations. Called actor-network theory (ANT)and less a theory and more a methodology, ANT, among other things, considers the objects of study to be on the same ontological level as the subjects, the humans interacting with the objects (Latour 2005, pp. 63–86). In science and technology studies this represents a radical shift in point-of-view (see Latour and Woolgar (1986)). For my purposes, this aspect of ANTis especially relevant when we consider objects, such as *syngvab*, which *can* act on their own and which do have at least presumed agency.

In addition, ANT challenges Western metaphysical assumptions about reality: that we assume reality is "out-there", independent of our actions, that it precedes us, that it is definite, and that it is the same everywhere Law 2004, pp. 23-26. Law shows in a number of examples that even if there is a reality "out-there", it is not independent of our actions (our measurement equipment, what he calls *inscription devices* influences the types of data we obtain), that it does not precede us (knowledge about a transcription factor only exists *after* we have discovered it), that it is not definite (forms are fluid depending on points of view), and finally, that it is multiple (different accounts of the same event can exist at the same time).

Important to ANTis the adherence to symmetry: "that we shouldn't let our ideas about what is true or false (in science or anywhere else) affect how we look at our subjects" (Law 2004, p. 40). This leads to the conclusion that we cannot build assumptions about our methods into our methods lest we simply verify the assumed methods in the first place. Thus, perhaps we should study humans and objects in similar ways, viewing that they give rise to similar abilities, such as agency. This is a sticking point for many, as it is certainly a challenge (as we saw in the previous chapter) for adults to ascribe agency to inanimate objects. Andrew Pickering, in his own take on ANT, writes that the "trajectories of emergence of human and material agency are constitutively enmeshed in practice by means of a dialectic of resistance and accommodation" (Pickering 1993, p. 567). This adherence to symmetric analysis is part of his description of the "mangle of practice", or the view that human and material agencies are "reciprocally engaged in the play of resistance and accommodation" (Pickering 1993, p. 567). The mangle, then, is the way in which the human and material "realms" are "interactively restructured with respect to each other"—and thus always studied in relationship to each other, symmetrically (Pickering 1993, p. 585).

With ANTand its symmetric analysis as a guide, we look for assemblages, for multiple realities, for creations of new social groups by the actors in the situation. We observe how new objects create new realities for the user, how, in the case of reflective HCI and, hopefully, *syngvab*, the object enables the expression of alternative forms of reflection. We look for situations where the object has left traces of its influence on the person.

⁂

We have a new way of considering the development of technology. We can now think of technological objects as more than simply instrumental devices. And we can study the use of these devices in a way that raises the object to the same ontological level as the human subject. New. New. New. Yet we now return to the old, to our drawing from a folk tradition for the design of *syngvab*, that of puppetry and the marionette. Along the way we will see how the use of these practices informs the development of new technological creatures and how we might be able to bring the puppet avant-garde to the design of novel robotics.

5

Puppetry and Performing Objects

> Does the idea of doing with art more than art still exist? Are the arts interested in more than themselves? Can puppet theatre be more than puppet theatre by giving purpose and aggressivity back to the arts and make the gods' voices yell as loud as they should yell?
>
> Schumann (1991, p. 83)

Up to this point in the thesis I have not written much of the creature first described in the introduction, *syngvab*. I first described it there as a "robotic marionette", a creature designed as a non-anthropomorphic, non-zoomorphic object with its strings pulled by computer-controlled motors. The questioning reader might wonder why I chose puppetry, and specifically marionettes, to draw from in the modeling of *syngvab*. In this chapter I will hope to convince you of why considering marionettes in particular is an appealing thing to do for new robotic creations. I will first outline some of the history of marionettes, specifically how it relates to folk performance and storytelling traditions. This will lead into how puppetry and marionettes has a strange relationship to robotics, not only through the desires of many roboticists, but also through the early text of Heinrich von Kleist (1972), "On the Marionette Theater", which suggested the possibility of an autonomous marionette. This will lead to a discussion of the small, but intriguing, set of extant contemporary robotic marionette and puppetry projects

Figure 5.1: Extant examples of marionettes. Image on the left is from `http://flickr.com/photos/praatafrikaans/255172011/` while the image on the right is from `http://flickr.com/photos/eugene/18142113/`.

5.1 PUPPET STUDIES, PERFORMING OBJECTS

There is an unfortunate dearth of material on puppets and their place in society. Beyond their contemporary all-too-common position as merely objects for entertainment, or, in the United States, for children, how do puppets reflect (or refute) the common themes of the day? How might puppets be related to broader concerns in the dramatic arts? In what areas of dramatic performance might puppets suggest new ways of thinking? And how can we manage the difficulty (and consequent possibilities) presented by the juxtaposition of an old, primarily folk art form with present-day technical options? This lack is all the more urgent when, in my opinion, the possibilities are great for drawing from this tradition and applying it to current concerns.

A dearth does not mean a complete absence, and there are fortunately a number of excellent articles and monographs that do deal with these concerns. In many cases, however, the texts that involve puppets do so in an ancillary way, focusing instead on broader themes of ritual, art history, semiotics, and cultural development (Bell 2001*a*, p. 5). Outside of academia, all too often authors (and many in the general public) have relegated puppets to the realm of curiosity, or as "merely" part of long-past, outmoded folk traditions.

Puppets can be seen as belonging to a broader realm of *performing objects*, "material images of humans, animals, or spirits that are created, displayed, and manipulated in narrative or dramatic performance" (Proschan 1983, p. 4)[1]. The performing object is often a part of extensive and well-developed folk traditions, as in Russia and the (present-day) Czech Republic (Bogatyrev 2001 [1923]; Reeder 1989). These traditions focus on story-telling and the ability of the puppet, as object, to stand-in for both characters of fantasy, as well as living figures for the purpose of satire.

In the United States there as not been such a tradition of puppetry. Indeed, up until the 1960s puppetry was left by the wayside, unconsidered by the broader public. However, since the 1960s there has been a strong development of puppetry that can be said to follow two divergent strands. The first is the work of Jim Henson and his Muppets. While indeed aspects of his *Muppet Show* could be said to appeal, on the subtextual level, to adults, the majority of his work is focused on children, especially through *Sesame Street*. The popularity of Henson's take on puppetry, while enhancing the image of puppetry in the US as a whole, unfortunately leads to the commonly-held belief that puppets and marionettes are child's play, something not to be concerned with as an adult.

The other strand of puppetry development, however, refutes this position. Peter Schumann and his Bread and Puppet Theater have, since the early 1970s, produced a number of large-scale works, most notably *Our Domestic Resurrection Circus*, that address political and social concerns through the medium of puppetry (Bell 2001*b*). These shows are decidedly for the adult, using the ability of the puppeteer(s) and her puppet by "using the automatically evocative abstract symbolism of puppets and masks" to focus on the historical and contemporary successes and problems of the United States (Bell 2001*b*, p. 54). For Schumann, and in direct relationship to our concerns about the voice discussed in Chapter 2,

> The radicality of the puppet theatre includes a redefinition of language as not merely a tool of convenient communication. Puppet language is more than an instrument of fine-tuned information. It is an experiment which strips words and sentences of their secondary fashionable contexts and condenses quantities of habitual gossip into singular terms. The puppets need silence, and their silences are an outspoken part of their language (Schumann 1991, p. 77).

While *syngvab* is not political in the sense of addressing political topics[2], it *is* political and radical in the sense of using the object, the strange form of the creature

[1] We could say, in the case of *syngvab*, that according to this definition it is *not* a performing object, as it is not a representation or "image" of a human, animal, or spirit. Yet, *syngvab* does participate in performance, and I would argue that even though the form attempts to be non-representational (see Chapter 6), the object is certainly a "material image" of some unknown *creature*.

[2] But see *policrae and demochi* in Chapter 9 for an example of a puppet-like creature that *does* address political issues.

and its actualization as a marionette, to encourage people to perform, to express themselves in a radical way.

5.2 WHY CONSIDER PUPPETRY? AND HOW DOES IT RELATE TO ROBOTICS?

So puppetry appears to be interesting and potentially relevant for considerations of radical transmission of new or subversive ideas. But why would we want to consider it in relationship to technological development? What could an old form tell us about how to design things today? And how might the puppet be related to the robot?

Conceptually, robotics is extraordinarily similar to puppetry, and specifically, the automaton. Indeed, the roboticist is trying to create the "intelligent machine", an puppet that comes to life and lives independently of the puppeteer (Sussman 2001). A roboticist tries to make an object, a creature, move in some believable way in order to extend or create a narrative, even if that narrative is simply daily life. The designer of a robot must understand how the movement of appendages, of arms, legs, the eyes and the head, work to support (or detract) from the desired effect. Like the puppeteer, a robotics programmer desires and requires a certain level of "disbelief" on the part of the observer, an agreement to partake in the perpetuation of something that isn't entirely "real".

The field of robotics would do well to consider this legacy of automatons and puppets, as the puppeteer knows much about creating a strong performance. As we will see in the next section, using the physical affordances of the marionette solves a number of difficult kinematic issues that the robotics field worries about. But more importantly, strengthening the link between robotics and puppetry ensures that roboticists and puppeteers alike understand the connection between the two fields, separated unfortunately by technology, but connected by a deep desire to solidify real-life through the creation of abstracting creatures.

5.3 VON KLEIST AND THE FUTURE OF MARIONETTES

There is a strange way in which a text written in 1810 presaged the present development of robotics and automata. Heinrich von Kleist, in his "Über das Marionetten Theater" ("On the Marionette Theatre") writes of his interactions with Herr C., a friend of his who is a dancer in a local theater. This text, deceptively short and written in a decidedly conversational style, has been extensively analyzed, especially in the psychoanalytic realm (Schaefer 1978). Without going into a deep space of analysis we can look at the text in another way: a predictive description of the potential power of an automatically controlled marionette or puppet.

Heinrich von Kleist, playing the role of the narrator in the tale, speaks with his friend Herr C. about some of the deeper questions of the marionette theater. von Kleist had come to the discussion with the view, soon to be found to be wrong, that the work of the puppeteer was "something rather dull: somewhat like grinding the handle of a hurdy-gurdy" (Kleist 1972, p. 23). Instead, Herr C. finds that the movement of the marionette was "exceedingly graceful in the dances" seen in the market (Kleist 1972, p. 22), something with which von Kleist, in the role of the narrator, concedes. Herr C.'s observations of the movements of the marionettes reflects well some of my reasons for choosing the form of the marionette for *syngvab*. Just like I will describe in Chapter 6, making the creature be a marionette enables me to use the "power" of physics for the purposes of movement. Herr C. mentions that von Kleist "must not suppose that every single limb, during the various movements of the dance, was placed and controlled by the puppeteer" (Kleist 1972, p. 22). Indeed, the movements of individual segments of a puppet or marionette are not individually articulated; there is often but one string controlling the distal end of a limb, the rest of the joints being left to move under the influence of gravity. This enables the puppeteer to be concerned less with the movement of each part and rather with a global curve or direction, letting gravity decide through its own laws the motion of the rest of the joints. Herr C.'s view is that this simple line of movement of the puppeteer is "nothing other than the path to the soul of the dancer" (Kleist 1972, p. 23), the "dancer" representing the connection between puppeteer and Herr C.'s occupation.

Herr C. saw that the line of movement caused by the motions of the puppeteer reflected a "trace of the intellect". Astonishingly, just like the sounds of the mechanical hurdy-gurdy, Herr C. believed that this "could eventually be removed from the marionettes, so that their dance could pass entirely over into the world of the mechanical and be operated by means of a handle" (Kleist 1972, p. 23). Similar to the views of many contemporary technologists, Herr C. thought that the development of an mechanical marionette would be in "such a realm only a God could measure up to this matter", meaning that "it would be almost impossible for a man to attain even an approximation of a mechanical being" (Kleist 1972, p. 24). Herr C. supposed that the movements of the mechanical marionette would not be subject to the laws of gravity, like the dancer is: the dancer must rest by coming close to the earth, while the marionette, especially in a mechanical version, would be able to break free from (certain) physical laws. Lacking a soul or spirit, the mechanical marionette, controlled only by handle or other physical system, would not have the trace of humanness that marks us as separate from a God. Thus, in Herr C.'s view, the mechanical marionette, due to its reliance on the laws of gravity, would be closer to perfection than a human puppeteer or dancer ever could be.

I do not share with Herr C. the view that a mechanical marionette, such as what I have developed with *syngvab*, is closer to any particular "perfection" or "ideal". Yet it is interesting indeed to note how Herr C. unwittingly predicted the development of robotic marionettes, and precisely for the reason why I have chosen such a form. The field of robotics is rife with projects and research attempting to create realistic movement in both physical and animated creatures. Forward and inverse kinematics, motion capture: all try to mathematically model different ways of forcing a joint to follow a specific trajectory in space. Each technique has its own difficulties and strengths; yet all suffer from the common problem of simulation: that we do not have perfect knowledge of physical systems and simply cannot know how to coerce a series of interconnected parts and joints to move in a desired path. While movement of our bodies, as an adult, is rather simple and intuitive, transferring this intuition (and the fact that we do not have inner knowledge of the underlying mechanisms that cause the movement) to a robotic creature is considerably problematic. Yet if we were to create an underarticulated system, where the number of degrees of freedom (DoF)of control is less than the total number of DoFin the system, we can crib gravity for free, letting its pull control the directions of the other parts and joints in the way most "natural" (natural because of our experience as embodied creatures in a world with certain physical laws). Instead of attempting to control each and every joint of a system, we can instead place only a subset under our control, but gain the power of expressive and natural movement. Simplicity, instead of a complicated set of questions; the power of embodiment and physicality.

5.4 ROBOTIC MARIONETTES

After the exposition of the previous sections, it is perhaps not surprising then to consider the various, but still somewhat limited, projects that involve robotic or digital marionettes and puppetry. As we saw in previous chapters, this desire to use digital technology as a means to "enhance" an existing physical system is a common trope in the development of novel media interfaces. Much like the goals of Herr C., the desires of many developers is to create systems that outperform, at at least match, the performances or abilities of a human user. Indeed, as we will see below, developers revel in the ability to make a robotically-controlled marionette "fly" or "float" (Chen, Xing, Tay, and Yeo 2005). Even with these lofty and somewhat traditional goals the projects I will describe present an interesting application of digital technology to the mutation and development of a rather traditional folk art-form.

CONTROLLING A REAL OR VIRTUAL PUPPET

One of the most direct applications of digital marionette or puppetry is in animation. Like I described earlier for robotics, articulating the movement of virtual creatures

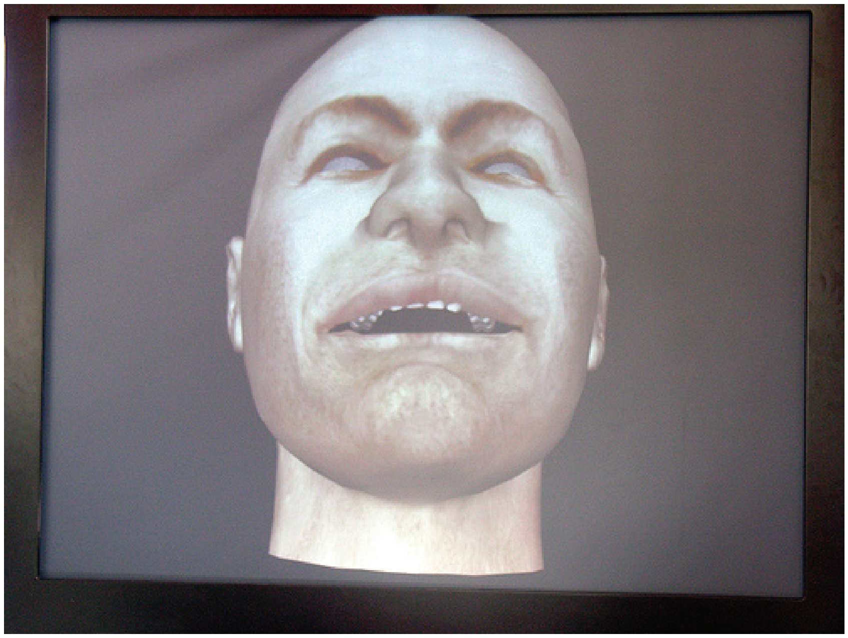

Figure 5.2: The Digital Marionette project by Corebounce. In the installation participants controlled the shape of the phase through marionette-like strings. Image from `http://flickr.com/photos/paolotonon/238248023/`.

for the purposes of animation is often dependent on a complicated combination of kinematics and specific, tedious placement of individual limbs or joints. Being able to simplify this process, while still retaining (or even gaining) the resemblance to physical motion, would be an enormous improvement. The work of Kim, Zhang, and Kim (2006) uses a haptic interface called a *Phantom*, a device that has a stylus that can be modeled to represent particular physical systems, to control the movement of an on-screen character. While the interface only allowed two degrees of freedom, this was still enough to have believable on-screen motion.

One can use the ideas of a physical marionette, such as the strings connected to a platform or bar, as the means of controlling an on-screen image. This is the tack used by Corebounce (2006) for their *Digital Marionette* project. Participants in their installation use a simple control mechanisms using rods and standard computer mice to control the expression of an on-screen character, again, like in Kim, Zhang, and Kim (2006), using a physical controller to manipulate a virtual character.

From another direction we can control a physical marionette with motion tracking data as in Yamane, Hodgins, and Brown (2003). In this work, the researchers asked an actor to move about in a manner that reflected the emotional content of a story; this was then mapped to the movement of a physical marionette. The results suggested that

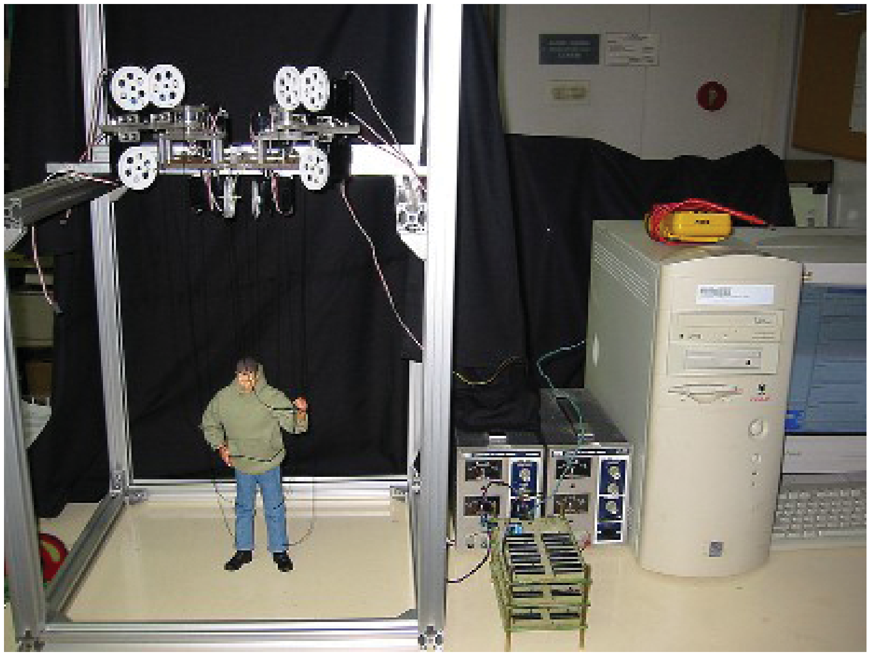

Figure 5.3: RObotic Marionette System (ROMS) of I-Ming Chen and colleagues. This is from the third iteration, ROMS-III. I drew inspiration for the motor system of *syngvab* from their work. Image from `http://155.69.254.10/users/risc/www/enter-intro.html`.

their system was able to translate the motions of the actor into believable movements on the physical marionette.

robotic marionettes in an ancillary position

One intriguing use of a robotic marionette is as a "capstone" design problem for college students (Wang, Liu, and Price 2006). This is due to the numerous complexities a physical robotic marionette system presents: the construction of a suitable "stage", the electromechanical system, and the software control of a many-degree of freedom system. Such a project, while not focusing on the marionette *per se*, affords students an opportunity to connect the old tradition of puppetry with modern technical abilities.

roms

Perhaps the most developed robotic marionette project is RObotic Marionette System (ROMS) by I-Ming Chen and colleagues at the Nanyang Technological University in Singapore (Chen, Tay, Xing, and Yeo 2004; Chen, Xing, Tay, and Yeo 2005). An image of one of their marionettes can be found in Figure 5.3. The design of their system draws from both Western and Chinese puppetry; the stationary attachment of motors and pulleys is reminiscent of the *Gou Pai*, where all the strings are attached to a single plate controller.

ROMS is capable of a number of highly articulated motions due to its high number of DoF. Additionally, the development of ROMS was influenced by discussions with a traditional marionette artist, helping to guide the creation of the system. Low- and high-level motor control enables the creation of gross and fine motion in the marionette.

❊ ❊ ❊

I have now detailed the conceptual underpinnings of the thesis: the use of non-speech sounds to enable the expression of the unspeakable; the strong ties that bind humans and objects together; the new ways in which we must conceptualize and study technology; and the strengths of considering puppetry in new robotic systems. It is now time to turn to the physical parts of the thesis, the development of the robotic marionette as object, and the software agent system.

6

Design of the Robotically-Controlled Marionette

Now that we have explored in some depth the theoretical underpinnings of this thesis, I would like to turn to the physical objects, namely the robotically-controlled marionette *syngvab* and its agent system. Throughout my description of the design and construction of these artifacts I will make note of where certain decisions reflect the concepts I discussed in previous chapters. Since a document such as a thesis is necessarily *linear*, the best I can do to make these connections is to briefly note them; however, you should see these attempts at tying long-distance (in text) threads together as really reflecting the interconnected nature of this thesis.

Where should we begin when speaking of the design of the marionette? Choosing a form for the creature is not as simple as making a variation on an existing animal or person; because of the design goals of the thesis this is simply not an option. I'll work through some of the resulting challenges that consequently inform the final design of *syngvab*. The current version of *syngvab* does not exist in isolation, of course, and I'll write of the earlier iterations, with some time spent on the first version, called *syngvaa*. Next, I will describe the stage for the creature, the environment in which it moves. Rather than being an afterthought, the stage and its design reflects considerations of performance and lighting that help to support my design goals. With the physical structure and the motivations for its design described I will then talk about the motor system: the structure of the motor control and the command-set used to control it.

Finally I will write of the simulator for *syngvab*, necessary to try out some of the agent control ideas that I will describe in the next chapter. This extended description of the physical form of the marionette (and its digital simulation) will leave us in a good position to explore the software agent behind the creature.

6.1 WHAT LIVES AROUND ANTHROPOMORPHISM AND ZOOMORPHISM IN FORM

In the earlier chapters I have written about theoretical "why's": why it is worthwhile to encourage people to explore non-speech vocal sounds, why people might form deep and meaningful relationships with objects, why I want to study people's interactions with those objects in non-standard ways, and why puppetry and marionettes are an appropriate tradition from which to draw. It is now time, however, to hack away at the details of my object, *syngvab* . I want to convince you, the reader, of my reasons for the admittedly strange form that *syngvab* developed. Rather than being ancillary points, the design of the creature *and* its surrounding environment reflect deep considerations as to our experiences with novel objects and ways to influence behavior.

No-one comes to new situations in life without prior experiences. Even infants have some traces of experience in the womb, not to mention limited knowledge of how to interact with their environment (although the standard "blank slate" view of sensorimotor abilities is being challenged in developmental psychology; see Bertenthal 1996 for a review). While it may not be in vogue to say something is *impossible*, it simply *is not possible* for someone to come to new events, new possibilities, with a blank mind and absent thoughts and memories that will undoubtedly influence their experience.

Similarly, no matter the desires of the designer, a new object is always situated not only in an environment that is often outside of her control, but also in relationship to a person's lifetime of experiences with other objects, also outside of the designer's purview. We simply cannot assume, as the designer is wont to do, that her creation will overthrow the prevailing world-view of the user, erasing traces of what had existed before. There is no global "New" command that a novel object magically possesses. Even if the designer attempts to create an entire *experience*, developing the environment and everything in it, the user still comes to the environment with an non-designed lifetime, and leaves the environment to return to a non-designed life. Lest I speak of an abstract designer, let me turn the spotlight directly on myself: I, as a designer of an original object and experimental experience, must infuse my process with these considerations. As designers, we should *embrace* the fact that people have other experiences, that we are not faced with the burden of designing everything. We can draw on what we know as observers of life to develop our objects, our environments, our buildings and experiences, in ways that perturb *in our own particular way* the lives of the users. This is not a handicap, but rather a possibility.

Since people will come to their interactions with *syngvab* endowed with their own prior experiences, the space of which I cannot predict or know of (nor would I want to), what are the consequences? I am faced with a seeming paradox: I want to create an original, experimental interchange between creature and person, even though I can never make something entirely original due to the variety of people's own lives. Yet I can use this paradox as an indicator for my design. Instead of fighting the battle between being original and derivative of what exists already, I can try something different: I can see what might exist in the negative space *around* those oppositional poles. What might we come up with if we looked at what *does not* exist around those poles, what is left sitting at the base or is off in the distance? What if we were to use this to our advantage?

CHALLENGES IN THE DESIGN OF *SYNGVAB*

Before I go into depth and write of how I drew from these advantages in the design of *syngvab* , I should first explain why *syngvab* provokes such challenges. While I have abstractly written of these issues in the earlier theoretical chapters, I want to collect and expand on them here.

THE PROBLEM OF CREATING NON-SPEECH SOUNDS

Fundamentally I am asking people to do a rather strange thing: make non-speech sounds. While someone who has experience in performance art, or in drama, or with working with the developmentally disabled might not find the idea of making sounds outside of language to be especially challenging, for many people it is. There are no experiences I can think of in our daily lives where we find ourselves producing sounds that are not phonetic. Language codifies the range of sounds that come out of our mouths, and while there are of course variations between people, the quantized repertoire is countably limited. Yet what happens when I *specifically* ask someone to move beyond the natural limits with their voice? How, as a designer, do I create a situation that enables this to occur? And what parts of people's lives might I be able to draw upon to simplify my task?

THE SITUATION OF MAKING THESE SOUNDS

So we know that what I am asking people to do is already difficult. How do I go about making this situation? And where is it appropriate for these sounds to be made? I have to create (or usurp) a place that is not threatening for the user. My peculiar request must take place in a decidedly familiar location and occur through an inviting situation.

FORM OF THE CREATURE

Strange request to make sounds. Inviting situation in which to make these sounds. What should the creature look like? There is a delicate balance to achieve. On the one hand, *syngvab* should not be too similar to existing animals, people or not. I have to fight the prevailing trend to make robotic creatures that resemble existing ones or that have components, such as legs, arms, fingers and toes, that are found in a variety of existing species. Similarly, I have to fight the opposite trend to make a robot that is simply a geometric shape such as a circle or a sphere. Why my strong opposition to these trends? Because the form of *syngvab* should reflect the way it is going to move in response to non-speech sounds. I can ask myself: how would a teddy-bear move to a cross between a shout and a cry? How would a box act to the sounds of a low rumble in the throat? I have no idea. I can figure out, nonetheless, how a relatively novel creature might do so.

On the other hand, we need to keep in mind the points that I made at the start of this section. People will never come to their interactions with *syngvab* with a blank slate, with no reference to existing animals and creatures. I need to design something that is familiar enough to not be immediately off-putting. The form cannot demand an inordinate amount of attention and thought. Partially this can be addressed through a relatively novel form for the creature, as well as the embedding of the creature in a familiar, inviting situation.

HOW A CREATURE RESPONDS TO THE SOUNDS

Finally, the creature must respond to the sounds in a way that, at least at first, makes sense. I cannot simply map sound ⇒ movement; to do so would immediately lose the attention of the user. Rather, the creature, in all its strange design yet familiar situation, must move in some understandable way. The situation does not have to remain this way, however. After a certain period of time with the creature, after the person has gotten to know the idiosyncrasies of the creature's movements, I can perturb the interaction in a way that is puzzling and forces the user to make sense of what he sees.

⁂ ⁂ ⁂

These challenges are indeed rather daunting. At a basic level I am trying to create the impossible: a novel creature that references earlier experiences but is not limited to them, in a familiar situation but with an original request.

It would be beneficial at this point in the narrative to return to the first iteration of this project, *syngvaa*. Some of the design choices made in this early creature can be seen, in modified form, in the design of *syngvab*. Nevertheless, an detour in time, back to my early desires for the project, will show how the current creature and environment better reflects the challenges and desires of the project.

6.2 THE DESIGN OF *SYNGVAA*

The original motivation for this series of creatures was to study how people interact with computational objects, those objects in our world that are infused with "presumed" intelligence, often through the use of artificial intelligence algorithms. My thought was that through an exploration of somewhat bizarre interactions between person and object, especially when that object could do things on its own, we would get a better understanding of how people might react to these object in more prosaic situations. By focusing on the relationship that would form as a result of intense, long-term interactions we could learn not only about a single person's idiosyncratic response to the situation, but also about what things designers of these systems and objects should keep in mind when trying to create novel objects and situations.

The first object of this thesis, *syngvaa*[1], was an attempt to study these questions through the construction of a mobile, embodied creature that would provoke people to respond to it in a way that blurred these boundaries. The creature moved in response to singing, or sounds that were like singing. *syngvaa* was quite simple in its construction. I designed the shape such that it would not directly resemble humans or animals. Of course, as I described earlier in the chapter, it is impossible for me to create a design that is entirely a blank slate; however, I can do what is possible to move the outline into a space that is far away from existing people or animals.

Movement came from two microcontroller-controlled DC servo motors and a rear castor. Analysis of the voice came from a separate computer with a basic pure data (PD) patch that analyzed the pitch and amplitude content of the incoming vocal signal. Commands to move the creature were sent via wireless RF transmission from the analysis computer to the microcontroller on the creature. Images of *syngvaa* can be found in Figure 6.1. The connection between sound and movement was quite unsophisticated: a change in pitch upward would cause it to move forward and vice versa. As amplitude climbed above a certain threshold, the movement would change from a straight line to a curved one.

[1] A note on naming: The word "syngva" comes from Old Norse meaning simply "to sing". Rather than enumerating each new revision of the creature with the suffixes "Version 1.0", "Version 2.0", and so on, I have decided to add letters to the end of the word instead, moving to the next letter of the alphabet with each new revision. Thus the first version is "syngvaa", the second is "syngvab", and so on.

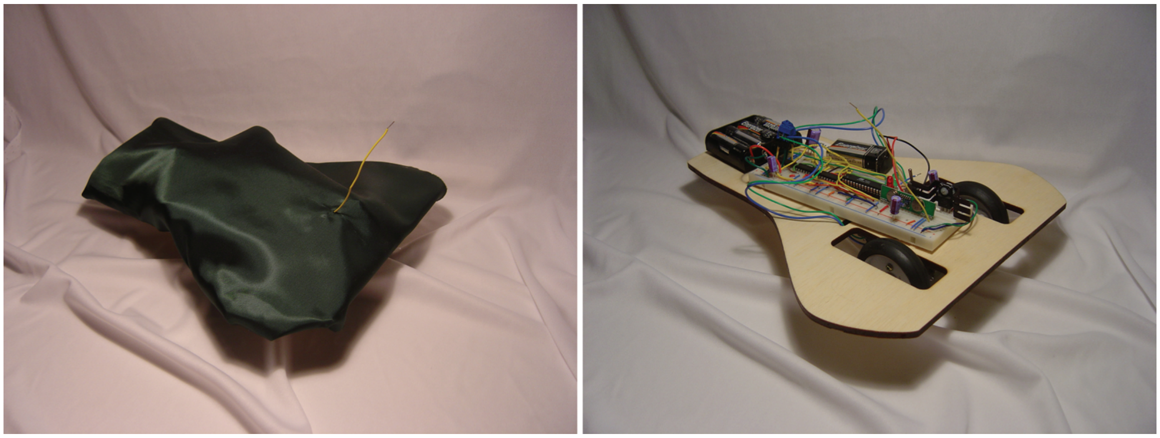

Figure 6.1: Original design of *syngvaa*. *Left*: Creature with cover. *Right*: Creature with cover removed, with internal electronics visible.

One thing to note about *syngvaa* was that its basic construction meant that it had no obstacle avoidance; it had deficiencies in wireless transmission because of the weak radio link; and the analysis of the voice was often poor due to lack of tracking of silences. This meant, however, that I (as well as other users) were *forced* into working through these problems via our behaviors. For example, consider the series of images in Figure 6.2 that come from a video of interactions with *syngvaa*. As *syngvaa* would get caught in corners, I had to change the way I was singing in order to prevent it from remaining stuck. There was no subsystem for finding the edge of the room; I, as the human in the interaction, knew where the edge was and thus I had to figure out the right sounds to make to extricate the creature from the situation. Additionally, *syngvaa* would sometimes move on its own. Because of the half-duplex nature of my wireless transmission system, the controlling computer had no way of knowing whether or not *syngvaa* had received a command. Thus, the program would flood the channel with a set of the same packets in the hope that at least one would reach the creature successfully. This rudimentary protocol (if it could even be called that) caused *syngvaa* to move unpredictably, or to repeat the same action multiple times. In turn, I responded to its movements "on its own" with my own sounds, creating something of a dialogue between myself and the creature.

In observations with a small set of users of *syngvaa* I saw a number of surprising things. First, in regards to its shape, people often asked me what it was "supposed to be". After I told them that it was not designed to resemble anything, they would offer their own interpretations: a lobster, a "creature with a tail". Because of the way I placed the wheels and the castor, people would often disagree as to what was the front or the back, no matter how the creature ended up moving. Finally, I observed that

Figure 6.2: Video stills for early movements with *syngvaa*. Video available from `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/syngvaa.mov`.

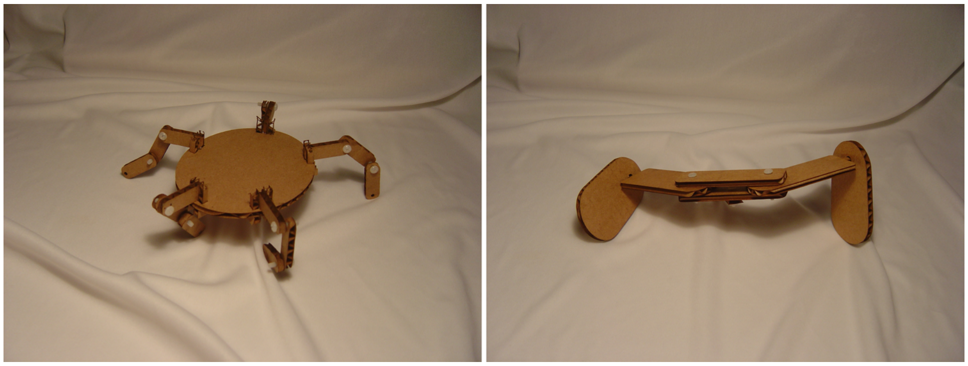

Figure 6.3: Some of the design candidates for the new version of *syngvab*.

the act of asking someone to sing, *for the purposes of making the creature move*, acted as a social catalyst. Staid businessmen would make all manner of sounds around their colleagues, oblivious to the normal social rules that prevent such actions. They became entranced in seeing what *syngvaa* would do in response to their voice.

NEW CREATURE DESIGNS

Following the construction of *syngvaa*, I wanted to consider other options for mobile creatures that would move in interesting ways. Early designs were again "autonomous" in the sense that they were supposed to move about on the ground without wires and without constraints. Some examples can be seen in Figure 6.3.

As I worked through possible designs for the successor to *syngvaa*, I considered a number of mobile options that would move about the floor: balls with offset masses, hand-like creatures, vehicles with offset wheels (Figure 6.3). Each prototype suggested a unique form of movement and had they been developed to fruition would likely have been engaging in interaction. However, all of the design suffered from technical considerations that were, due to design constraints, were show stoppers:

- *Need for batteries*: Each mobile creature would require a regular power source detached from any wires. This necessity for batteries would have imposed severe constraints on the amount of time the creature could run before needing to be recharged. As well, any requirement for regular charging of the creature by a user would have created an impediment and dependence on the user that I did not want to impose—the creature should be able to be mostly self-sufficient, freeing the user from the caretaker role.
- *Need for obstacle avoidance*: If the creature were to move autonomously about the floor it would need to have some knowledge of what is an obstacle. It would

need to know how to avoid them, how to remove itself from difficult situations, and so on. Roboticists have been writing obstacle avoidance algorithms for as long as there have been robots, and there is still no one general algorithm that works in a general number of cases, without the need for regular human attention—especially in a "noisy" environment like a home. As I want to reduce the human caretaker burden as much as possible, a marginally-functioning obstacle avoidance system was not an option.

6.3 THE FORM OF *SYNGVAB*

Thus I came to a sticking point in the development of *syngvab*. However, once I realized the power of puppetry (Chapter 5) I could see that a stationary creature, perhaps controlled by wires like a marionette, might be just as engaging as a mobile creature on the floor. I began to consider possible *syngvab* marionettes, one of which can be seen in Figure 6.4. This first marionette shows a strong resemblance to the original outline for *syngvaa*, reflecting the connection between the two iterations.

As I will describe in the next section, I at the same time became interested in the casting of light and shadow and the possibility of using this as an additional aspect of *syngvab* design. Since the creature would be stationary I would be able to easily control the location of light sources, both spot and diffuse. This suggested an open frame for *syngvab*, some early examples of which can be seen in Figure 6.5.

Given these early tests, I came up with the current iteration of the form of *syngvab* that can be seen, *in-situ*, in Figure 6.7. The next series of images show the following aspects of *syngvab* and its environment:

- The unfolded outline of *syngvab*, seen in Figure 6.6. This is done to enable the entire shape to be cut from a single planar piece of stainless steel using a waterjet.
- *syngvab* in the stage, with lighting, seen in Figure 6.8. The stage is approximately two feet tall, by two feet deep, by a foot and a half wide. Detailed construction diagrams can be found in Appendix B.
- Description of the parts of *syngvab* on the stage, seen in Figure 6.9.

6.4 A STAGE FOR THE MARIONETTE

The stage for *syngvab* is not merely a backdrop for the creature's movement, or merely an enclosure. Rather the stage is an embodiment of the performative metaphor that is

Figure 6.4: The first marionette version of *syngvab*. Notice the similarity of the body with the design of *syngvaa*.

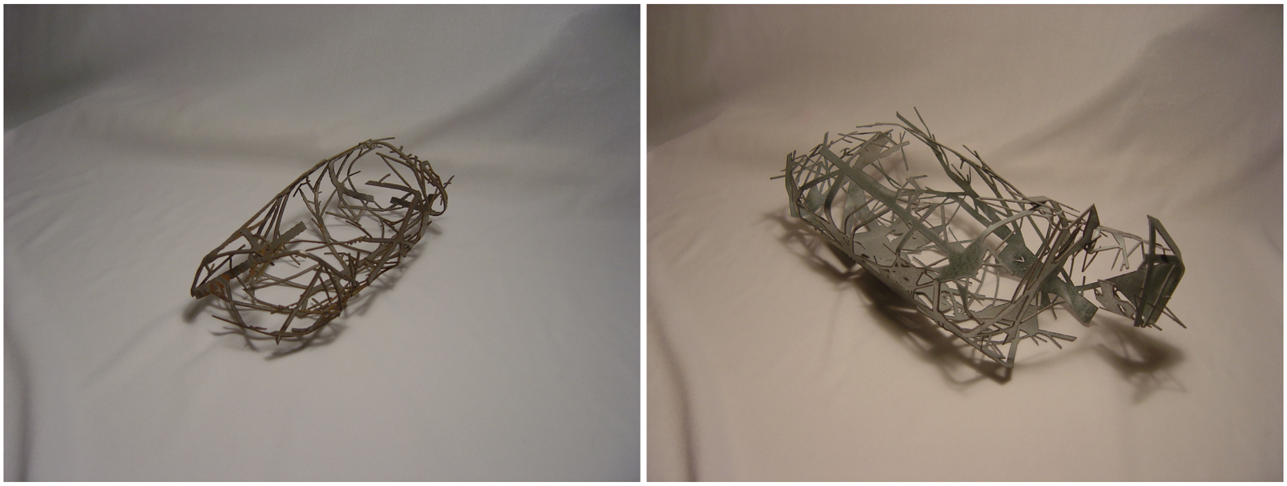

Figure 6.5: Candidate wireframes of the developed version of *syngvab*.

a strong conceptual component of this thesis. The stage solved a number of technical problems that I faced in the development of *syngvaa*, as previously described.

CONSIDERATIONS OF ABSTRACTION, LIGHT, AND SHADOW

If I were to design and build a robotically-controlled marionette that was to embrace the desire to dally around the poles of anthropomorphism and zoomorphism my first direction was to abstract things as much as possible. What might a marionette be like if it were made up of only sticks? Or planes? How could those geometric objects move in ways that were engaging for the viewer? Or that gave the perception of agency? This is not simply an academic or conceptual question. Psychological studies, beginning with the classic work of Heider and Simmel Heider and Simmel (1944), have shown that adults (and children (Ackermann 1991; Scholl and Tremoulet 2000)) readily interpret the movement of abstract shapes in terms of living actors, *i.e.*, people. Thus, it is reasonable to suggest that if I were to design an entirely abstract robotic marionette and carefully control its movement, people might respond to it as if it were a person, or at least were controlled by a person.

Additionally, the decision to create a marionette, and the subsequent realization of a need for a stage, enabled me to play with ideas of light and shadow. Various puppetry traditions work through shadow play, using the casting of light on screens as not only characters in the story, but also as ways to smooth transitions between disparate parts of a narrative. Unlike an autonomous creature that moves about the floor, with a marionette on a stage I can control, within a wide gamut of possibilities, the lighting of the scene. With appropriate backdrop and floor materials (that is,

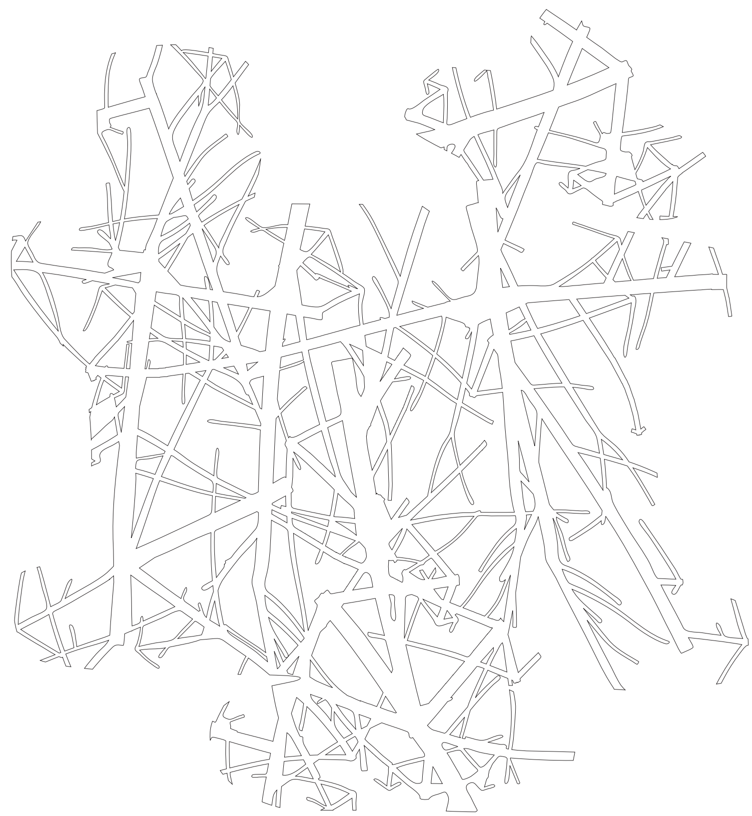

Figure 6.6: Outline of *syngvab* unfolded. The back of the creature is at the bottom, and the bottom of the creature is in the middle of the image. I created the outline by overlapping "trees" made by a generative process.

Figure 6.7: Closeup of *syngvab in-situ*, on the stage.

Figure 6.8: *syngvab in-situ* in a participant's home.

Figure 6.9: Diagram of *syngvab*, including the creature itself, location of motors, and stage spotlights. A video of the first movements of *syngvab* can be found at `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/syngvabFirstVideo.mov`, while a video of the final movements with the voice can be found at `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/syngvabVideo.mov`.

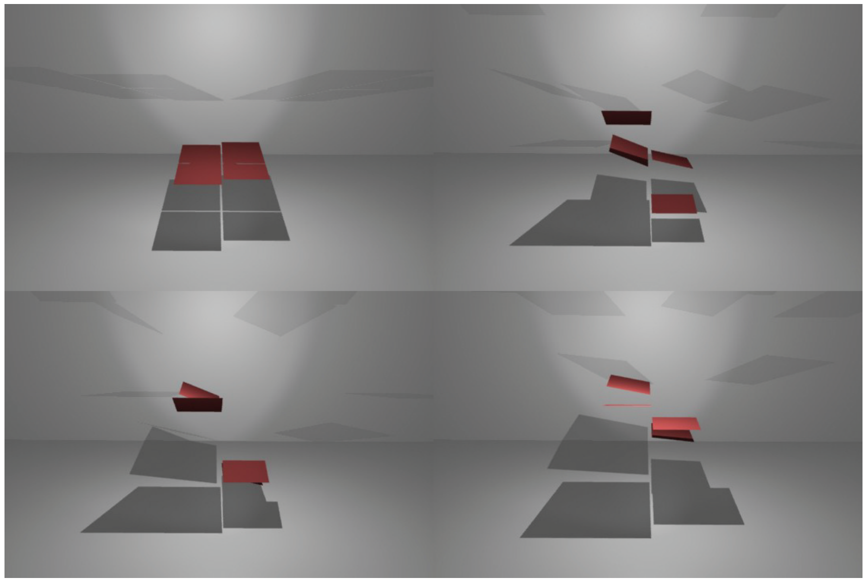

Figure 6.10: Screenshots of a modeling test using moving planes and lighting. Time moves from left to right, top to bottom. Notice how the pattern of shadows changes through time. See `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/planesTest.mov` for the animation.

materials that will show shadows well), the marionette becomes multiplied and transformed into a range of projected images, each one a possible focus of attention.

To study this further I used the open-source 3D modeling and animation software *blender* Foundation (2007) to mock up two types of abstract marionettes. The first was a series of four square planes, with "strings" attached to the four corners of each plane. There was one overhead diffuse light, and two spotlights at the lower-left and lower-right of the scene. Figure 6.10 shows a series of screenshots from one animation of this scene. We can see that because of the location of the lights, what is, in "actuality", a square plane becomes an extended trapezoidal figure when its shadow is projected on the floor or the wall. The effect is especially interesting when viewed in the context of the animation: as the planes move apart, the shadows become disjointed, separate and flying off in their own direction; yet as the planes come together again, the shadows do as well, with a certain perceptual lag that lengthens the tension.

To more faithfully model this situation I created another animation that incorporated a modeled stage, this time using a series of rods as the marionette (Figure 6.11). Additionally, I made one of the rods a curved bar, made from an extruded, smoothed Bezier curve. Lighting is as in the previous animation. Because of the

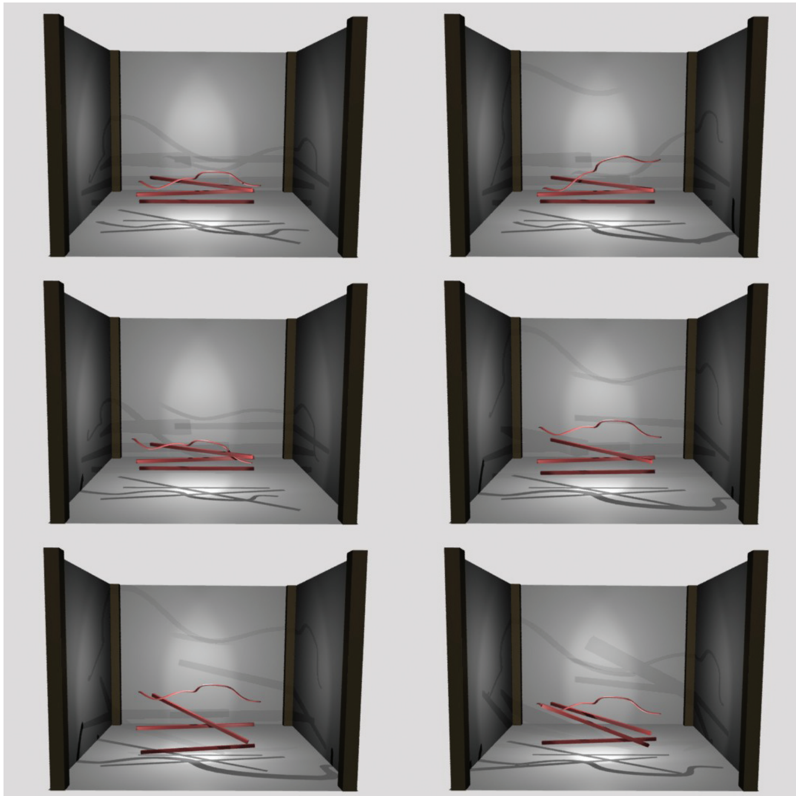

Figure 6.11: Screenshots of a modeling test using moving rods, stage, and lighting. Time moves from left to right, top to bottom. The pattern of shadows, because of the shapes of the rods as well as the stage, is even more intricate than in Figure 6.10. See `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/rodsTest.mov` for the animation.

starting orientation of the rods and the additional constraints due to the wall so the stage, the shadows cast by the rods are even more complex than in the case of the moving planes: the bent bar's shadow "jumps" along the left and right walls of the stage, the shadows of the straight rods intersect in unpredictable ways (at least for the casual observer).

Yet for all the engaging attributes of the lights, shadows, and abstract shapes, the designs of these "marionettes" was sadly lacking. Even if I could (potentially) attribute some type of intention to the geometries on the screen, their lack of resemblance to a *creature* made them unsatisfying. I worried about whether or not people would remain engaged over a long period of time with the shapes. Would someone care about a bar moving on the screen? Would they imagine what the *bar-as-creature* was like as an independent entity? Would they give it a name? My answer to all of the questions was no. I realized that even though I wanted to push as far away from anthropomorphism and zoomorphism as possible, I could never wholly break free—and that if I tried to, as in these shape studies, the results would be less than I desired. There had to be room for projection in the creature's design; there had to be some connection, however tenuous, with a person's image of the concept *creature*. As we saw in the earlier sections, this lead to the current design of *syngvab*.

The planes and rods studies, nevertheless, indicated the strength of using lighting as an additional component in the design of the stage of *syngvab*. In the current version of the creature + stage combination lighting does not serve to extend the range of the creature nor to suggest other attributes of its existence. Rather, the shadows cast by *syngvab* help to entice people, to encourage them to be interested in what is going on on the stage, to have them comment when the shadows (and thus *syngvab*) have done something unexpected. The lighting and shadows help complete the *mise en scéne* and further strengthen the appearance of a performance.

6.5 CONTROL OF THE STRINGS

I programmatically move *syngvab* through a set of motors connected to the strings of the creature. These motors are controlled by a custom motor-control system that enables me to easly connect new modules of motor control boards at will. This modular arrangement enables me to easily add or remove degrees of freedom for the creature as needed.

STEPPER MOTORS

DC motors are a type of electromagnetic motor where the shaft rotates when a voltage is applied to its terminals. Commonly the speed is controlled through Pulse-Width Modulation (PWM). Unless there is an external device, such as a quadrature encoder, it is extremely difficult to control (or know) precisely the position of the shaft. With stepper motors, however, this is not the case. Stepper motors work by moving in discrete, quantized "steps". In brief, the motor has a series of toothed electromagnets around the shaft. Each set of electromagnets is energized through a separate set of terminals. When each terminal set is energized in the correct sequence, the motor moves one step. Each motor has a different number of steps that completes one full revolution.

Stepper motors enable precise control of movement without the need for external parts such as encoders[2]. It is possible to have even finer control over movement

[2] This assumes there is no slippage of the motor. Slippage occurs when a torque is applied that is greater than the holding torque of the motor. For the creature in this thesis, slippage is unlikely to occur.

Table 6.1: Parameters of Stepper Motors Used

Applied Motion Products 5015-824	
Parameter	**Value**
Voltage (V)	8.56
Steps per revolution	400
Ohms (Ω)	26.5

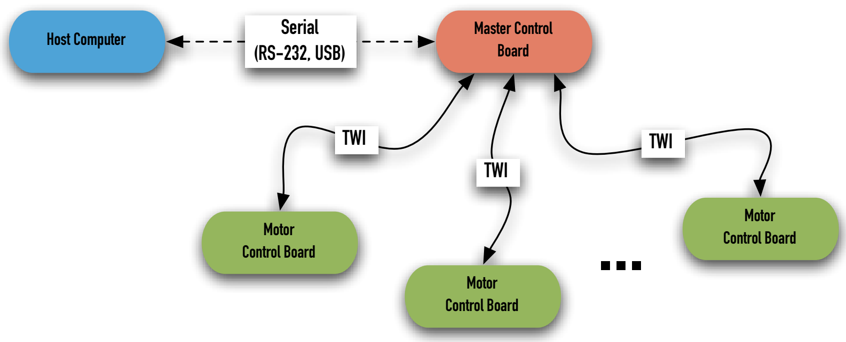

Figure 6.12: System design for the motor control boards

through the use of half-stepping or microstepping. In half-stepping, overlapping series of terminals are energized in sequence, allowing the shaft to be "in-between" two stable states. This doubles the effective number of steps for any particular motor. More esoteric schemes, such as microstepping, work by controlling the current that is applied to the motor; this requires a feedback circuit to monitor the motor.

Given that we want to have relatively definite knowledge of where the creature is at all times I chose to use stepper motors with half-stepping. The motors used, for all of the degrees of freedom, are from a surplus supplier[3]. The relevant parameters can be found in Table 6.5. For their price the accuracy and speed of the motors are more than adequate for our needs. I mounted the stepper motors to the ceiling of the stage using custom-designed brackets. The strings that are attached to the marionette are threaded through holes in the ceiling. On the shaft of each motor is a single pulley on which I attach the string. I chose the pulleys to reduce the number of revolutions needed to raise a given degree of freedom a certain amount, as well as to provide a groove for the thread to follow.

MOTOR CONTROL SYSTEM

The motor control board is a system of pluggable modules that communicate over a standard board-to-board protocol. The architecture of the system can be seen in Figure 6.12. There is a single master control board that speaks to a host computer over a standard RS-232 serial connection[4]. The master control board speaks to a daisy-chained set of motor control boards over the Two-Wire Interface (TWI), an implementation of the Inter-Integrated Circuit (I2C) protocol. Motor and logic power are also daisy-chained to reduce grounding issues as well as the need for a large number of separate power supplies.

MASTER CONTROL BOARD

The master control board provides coordination for the entire motor control system. A complete schematic can be found in Figure A.1. There is nothing that is remarkable about the hardware design. RS-232 serial communication and translation occurs through a MAX233 chip to reduce part count (as opposed to the more-common MAX232). Power is provided by a regulated 9-12V source, supplying 5V with a maximum of 5A output current. The motors I chose can be run at such a low voltage with little loss of torque. The core microcontroller is an Atmel ATMEGA168. See Figure A.1 for a schematic.

The firmware for the master control board, living on the ATMEGA168, provides a command-line interface that can be accessed through any terminal program capable of VT-100 emulation. The command-set is an extremely minimal, but complete, set of commands, as detailed in Table 6.5. Each command is terminated by a carraige-return, line-feed and there is a one-command history. Facilities for the command-line, including USART processing, command parsing, and command interpretations, are provided by the AVRLIB set of libraries written by Pascal Stang (Stang 2005).

MOTOR CONTROL BOARD

The motor control boards were designed to be a simple, pluggable set of modules that are controlled by the master control board. Each motor control board can control a maximum of four stepper motors through a single ATMEGA168 microcontroller. Stepper control comes from the ULN2065BN motor driver chip, a simple integrated circuit that can source up to 1.5A at 80V in a DIP package. Each motor control board speaks to the master control board over TWI. Power and ground for logic and

[3] Electronics Goldmine: `http://www.goldmine-elec.com/`.

[4] There is no reason why USB could not be used instead; RS-232 was chosen in this version of the board for expediency.

Table 6.2: Master Control Board Command Definitions

Command Name	Arguments	Description
`send`	`0xXX …`	Send a command to the currently selected motor control board. Arguments are given in hexidecimal format (0xXX). The first argument must be a command-code, the rest of the arguments are the necessary parameters for that command-code.
`receive`	`0xXX`	Receive information from the currently selected motor control board. The argument must be a valid request command-code. Only partially implemented.
`setBoardAddr`	`0xXX`	Set the currently selected motor control board to *0xXX*. The TWI specification allows a maximum of 127 devices to be connected to a single bus. In our design, this would allow a theoretical maximum of 508 motors to be addressed by a single master control board; however, computational considerations make this only theoretical, not practical.
`help`	`none`	Get brief information about each command.
`exit`	`none`	Quit the command-line interpreter; resetting the master control board restarts it.

motor supplies, separately, are daisy-chained amongst the master and motor control boards. (See Figure A.4 for a schematic.)

The firmware for each motor control board is where the bulk of computation takes place. We need a way to smoothly and easily move each motor a given number of steps. As well, we need a command-set that is general enough to allow us to provide necessary movement commands by an outside application, namely the *syngvab* agent.

CREATION OF MOTOR SPEED PROFILES

For smooth movement, it is not as simple as merely applying the necessary sequence of pulses to the motor for the desired duration. If we were to send 400 pulses to the motor we would see jerky movement: the motor would start from a dead stop, accelerate "instantaneously" to full speed, and at the end of the pulse sequence, deaccelerate "instantaneously" to a stop. Of course, this would not happen in zero time, like the word "instantaneously" implies, but the visual result is the same: object being moved by the motor would jerk at both the beginning and the end of the sequence. Depending on the inertia of the object and the number of steps this could last for the entirety of the desired movement.

This is not what we want. At the most basic level, jerkiness in movement is a hallmark of much of the general public's perception of robotics. Humanoid robots are betrayed, if not by their lack of emotion, by instead their lack of smooth movement. Since *syngvab* is to be put in people's homes, it would be desirable for it to move

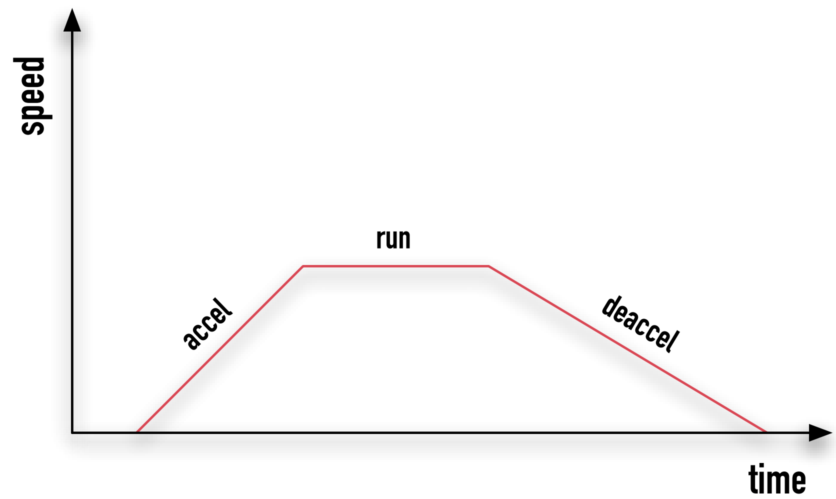

Figure 6.13: Motor profile graph, showing a generic trapezoidal profile with acceleration, run, and deacceleration phases.

in as smooth a manner as possible, limited by hardware and my technical abilities. Thus, I would like to create a *movement profile* that follows a trapezoidal outline: an acceleration phase, a constant run phase, and a deacceleration phase (see Figure 6.13).

However, the computational abilities of the motor control system is limited. The processor is an 8-bit microcontroller, ensuring that even basic math such as the square root is going to be taxing, no matter how fast we clock it. How do we deal with this situation? David Austin, in a recent article on *Embedded.com*, shows how to create arbitrary speed profiles in real-time on a microcontroller using a Taylor Series approximation to the square root (Austin 2004). The algorithm works by continually resetting the processor's timer counter to generate interrupts at, in the acceleration phase, shorter and shorter durations, and in the deacceleration phase, longer and longer durations. We then tell the motor to move one step during the interrupt callback.

Let $\delta t = \frac{c}{f}$ be the duration of the current timer interrupt, where c is the timer counter and f the frequence of the timer. Also let $\omega = \frac{\alpha f}{c}$ be the angular velocity, where α is the step angle and is defined to be $\alpha = 2\pi/steps\_per\_revolution$. We can write the complete movement of the shaft, in radians, between time 0 and time t as

$$\theta(t) = \int_0^t \omega(\tau) d\tau = \frac{\dot{\omega} t^2}{2} = n\alpha, \qquad (6.1)$$

where n is the total number of steps and $\dot{\omega}$ is the acceleration (or deacceleration). Thus, to find the duration of a sequence given a specific number of steps, we calculate

$$t_n = \sqrt{\frac{2n\alpha}{\dot{\omega}}}. \tag{6.2}$$

We can write the current timer counter value recursively as

$$c_n = f(t_{n+1} - t_n). \tag{6.3}$$

The initial timer counter is given by

$$c_0 = f\sqrt{\frac{2\alpha}{\dot{\omega}}} \tag{6.4}$$

Thus, in terms of the initial timer count c_0, the timer count for step n is

$$c_n = c_0(\sqrt{n+1} - \sqrt{n}) \tag{6.5}$$

Equation 6.5 needs to be updated constantly during the acceleration and deacceleration phases. Computation of two square roots will likely overwhelm the microprocessor and extend beyond the length of the interrupt. Thus, we need some way to simply Equation 6.5. If we take the Taylor series expansion of Equation 6.5 use only the first order terms we arrive at

$$\frac{c_n}{c_{n-1}} = \frac{4n-1}{4n+1}. \tag{6.6}$$

Simplifying further, to isolate the c_n term, we arrive at our final update equation of the timer count:

$$c_n = c_{n-1} - \frac{2c_{n-1}}{4n+1}. \tag{6.7}$$

Given the simplifications we made there will be some inaccuracies in the initial timer count. We can correct these through a slight modification to the calculation of c_0:

$$c_0 = 0.676f\sqrt{\frac{2n}{\dot{\omega}}}. \tag{6.8}$$

The Atmel application note 446 gives the details for implementation of these equations on the microcontroller and calculation of acceleration, run, and deacceleration profiles (Corporation 2006).

MOTOR COMMAND DEFINITIONS

The command set for the motor control boards can be found in Table 6.5. The commands are relatively self-explanatory. For each type of command I offer the ability to set it for all motors at one, a specific motor by itself, or for all motors in a SEQUENCE. Additionally, for the motor profile commands, we can use defaults set either at startup or by calling the DEFAULT commands. For 16-bit values I split them into two bytes, one for the high part, the other for the low part.

6.6 SIMULATION OF *SYNGVAB*

As we will see in the next chapter, when developing the movement of *syngvab* in response to sounds, I often will need to try out a large number of possibilities, some of which might be damaging to the creature. Of course I (or the genetic algorithm) will eventually eliminate those options, but it is better to come up with a situation that will not cause physical problems with the creature. I can do so through the development of a physical simulation that captures all of the relevant movement properties and enables, given the inherent limitations in simulations, the visual inspection of different types of movements.

While there are a number of different physical simulation libraries available, I chose to use the open-source Open Dynamics Engine (ODE) (Smith 2006; Smith 2007)[5]. ODE is for simulation of rigid-body dynamics using a first-order integrator. Its primitive types include boxes and cylinders, with a variety of joints such as ball, hinge, universal, and slider. I used the open-source PyODE library, which integrates ODE with the python language, and displayed the results using OpenGL (again, in python using the PyOpenGL bridge).

To simulate *syngvab* we decompose the creature into a series of boxes. The main body of the creature is modeled as a single box. I model the chains as a series of interlocked cylinders, with the attachment point to the creature a hinge joint and each joint between the individual cylinders a hinge joint as well. The "arms" of the creature are attached to the body of the creature through a universal joint to enable both up-down and left-right movement, much as in physical creature. Figure 6.14 shows the simulated version of *syngvab* in a variety of views, while Figure 6.15 details the joint connections.

As I will describe in more detail in the next chapter, Chapter 7, an output manager controls sending movement commands to both the physical creature and the

[5] While ODE is stable and well-developed, see also the new Bullet Physics Library (`http://www.continuousphysics.com/Bullet/`), which is integrated into the Blender (`http://blender.org`) 3D modeling program.

Table 6.3: Motor Command Definitions

Command Name	Value	Parameters	Description
		Global Commands	
EMERGENCY_STOP	0x00	None	Immediately turn off motor command processing. Reset the microcontroller to re-enable processing.
		Direction	
SET_ALL_DIRECTION	0x40	direction	Set the direction for all motors
SET_MOTOR_DIRECTION	0x41	motorID direction	Set the direction for the given motor ID
SET_MOTOR_SEQUENCE_ DIRECTION	0x42	direction1 direction2 direction3 direction4	Set the direction individually for each motor in the sequence
		Movement Profiles	
SET_ALL_PROFILE	0x60	stepsHigh stepsLow accelHigh accelLow deaccelHigh deaccelLow speedHigh speedLow	Set the trapezoidal speed profile for all motors connected to this board.
SET_MOTOR_PROFILE	0x61	motorID stepsHigh stepsLow accelHigh accelLow deaccelHigh deaccelLow speedHigh speedLow	Set the trapezoidal speed profile for the motor identified by *motorID*
SET_MOTOR_SEQUENCE_ PROFILE	0x62	stepsHigh stepsLow accelHigh accelLow deaccelHigh deaccelLow speedHigh speedLow (x4)	Send motor profiles for each motor in sequence.
SET_ALL_PROFILE_ DEFAULTS	0x63	stepsHigh stepsLow	Set the speed profile for all motors with the given steps using the default acceleration, deacceleration, and speed
SET_MOTOR_SEQUENCE_ PROFILE_DEFAULTS	0x64	stepsHigh stepsLow stepsHigh stepsLow stepsHigh stepsLow stepsHigh stepsLow	Set the speed profile for each motor for the given steps, using the default acceleration, deacceleration, and speed
SET_DEFAULT_ACCEL	0x90	accelHigh accelLow	Set the default acceleration
SET_DEFAULT_DEACCEL	0x91	deaccelHigh deaccelLow	Set the default deacceleration
SET_DEFAULT_SPEED	0x92	speedHigh speedLow	Set the default speed

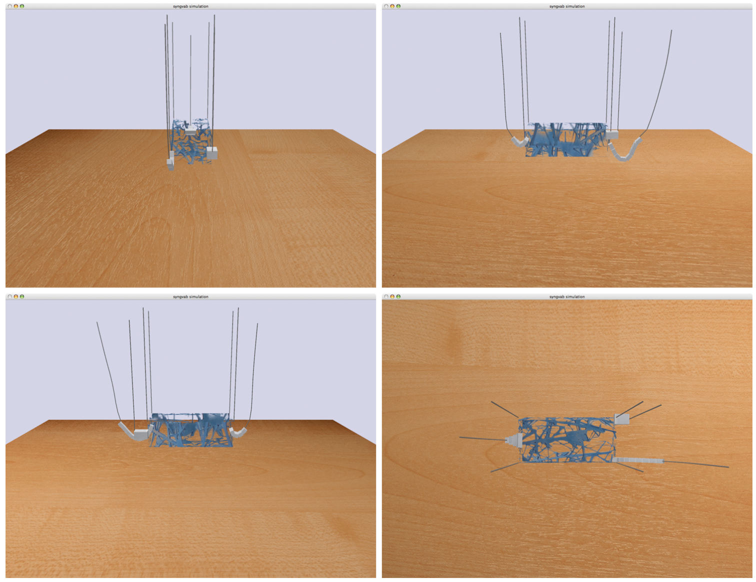

Figure 6.14: Views of the simulated version of *syngvab*. Clockwise, from upper-left: front view, left view, right view, top view. A fly-by video can be seen at `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/syngvabSimulatedFlyby.mov`, while a video of the simulated creature moving to simple motor commands can be seen at `http://purl.oclc.org/NET/NKNOUF/MSThesis/Movies/syngvabSimulatedMovement.mov`.

simulation. The simulator translates a series of motor commands, such as `SEND_MOTOR_SEQUENCE_PROFILE_DEFAULTS 0x01 0xff 0x01 0xff 0x01 0xff 0x01 0xff` into a set of durations for each of {acceleration, speed, deacceleration}.

One of the challenges is how to convert a given number of steps, at a given acceleration, into a desired force for a certain duration. In the simulation I go through the series of equations that occurs on the physical creature, previously described above, tallying the total duration spent in the acceleration, run, and deacceleration phases. These values then determine how long I apply a force to the simulated creature. I update the force by implementing a basic Proportional-Integral-Derivative controller (PID) controller. Current acceleration (deacceleration) comes from computing $\text{currentAccel} = \text{currentVelocity} - \text{previousVelocity}$ using the ODE command `getLinearVel()` on the final (end) body in each chain. This acceleration (deacceleration) is compared to the desired acceleration in the PID controller: $\text{desiredAccel} = \text{PID}(\text{currentAccel})$. The desired force is then $\text{desiredForce} = \text{holdingMass} * \text{desiredAccel}$. For each timestep in the simulation, until the desired acceleration (deacceleration) phase is over, I update applied force through these series

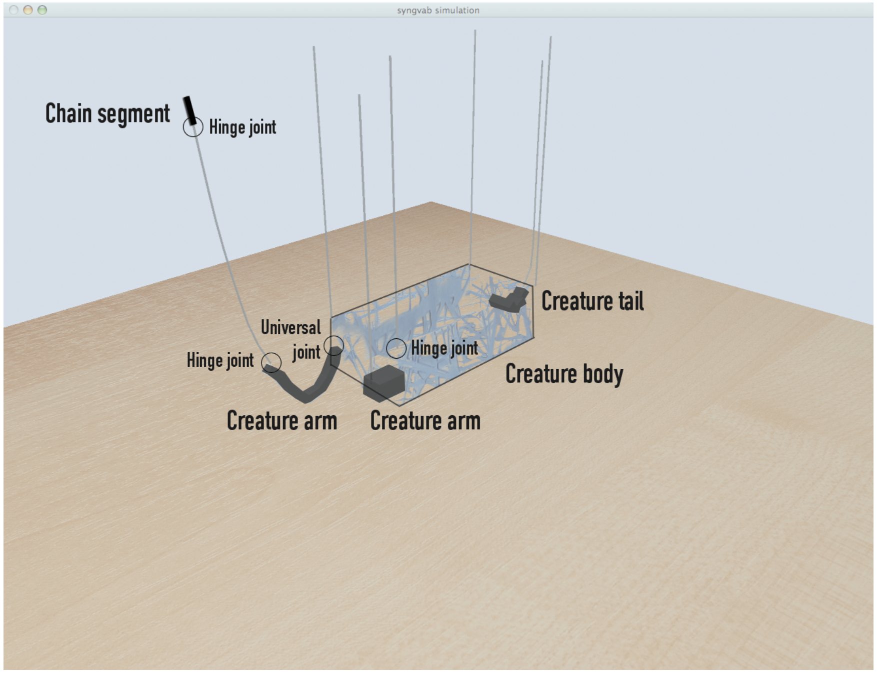

Figure 6.15: Diagram of simulated version of *syngvab*. Most of the joints in the creature are modeled as hinge joints, with the exception of the connections of the arms to the creature body, whcih are modeled as universal joints.

of equations. The run phase is similar, except I calculate the force needed to keep the chains moving at the desired *velocity* instead of acceleration.

These values are all summed with a *holding force*, or the force needed to make the creature "float" above the floor of the simulation. This is a simple calculation of the force needed to maintain a position in the presence of gravity. I should note that the mass values used in all of the calculations are not those of the actual creature, but were chosen to make calculation expedient and a *relatively* correct.

I send commands to the simulator through Open Sound Control (OSC) (Wright and Freed 1997; Wright, Freed, and Momeni 2003), a network protocol often used for sound and multimedia applications, but also general-enough for many types of low-bandwidth process-to-process, computer-to-computer communication.

SIMULATION AND PHYSICALITY

With the description of the simulated creature here, one might ask why I went through the difficulty of creating the physical object. Why not just have people interact with a screen-based system instead of spending the time making an embodied, physical thing? Much agent work is based on the belief underlying this question: that an on-screen agent will be just as engaging to interact with as one that is embodied. However, recent work by Cory Kidd and colleagues shows that this is not the case (Kidd 2003; Kidd and Breazeal 2004). The result is exactly the opposite: people found that an embodied robot is more engaging than one that is merely animated (Kidd and Breazeal 2004). This is perhaps intuitively understandable: interactions with a real thing will be more powerful than with one simulated on-screen. Thus I created the physical object of the thesis, *syngvab*, instead of merely simulating it.

⁂ ⁂ ⁂

Now that we have a physical object and a general control-system to make it move, how do we make the movements in response to non-speech sounds? But before that question, how *should* we make it move, and how do we determine whether or not a sound is speech or non-speech in the first place? The next chapter confronts these issues and describes the design and implementation of the agent for *syngvab*.

7

Design and Implementation of the Software Agent

An intricate, smoothly moving motor system is going to remain inert and nothing more than a piece of kinetic sculpture unless I create a system that makes the creature move in interesting ways. In this chapter I write of the agent of *syngvab*, the programming structures that enable the creature to move in response to non-speech vocal sounds. Along the way I describe what a (nearly)-ideal agent would be like, showing why such an agent cannot exist (yet). Given this impossibility, I describe the four components that make up the agent: Listen, Consider, Choose, and Do. Finally I write of how we might evaluate the agent in use, leading to the descriptions of interactions with *syngvab* in the home.

7.1 AGENT DESIGN

At some level it would be desirable, from both conceptual and philosophical standpoints, to design an agent such that it is *not* decomposable into individual parts. Why would this be? While we do not know with much certainty the structure of living systems we can safely assume that biological intelligent systems are not built from discrete modules with limited communication amongst them. Rather, current research suggests that representation and processing is distributed across many areas and scales, dismissing the notion that we can easily decompose the brain into coherent, single, units of understanding.

Yet when it comes time to develop an agent-based system we face the reality of modern, accessible programming languages (Downie 2005, pp. 152–161). Many, including the standard C, C++, Java, and Python, all still, at their root, procedural, sequential languages. Threading support, and the consequent ability to switch amongst tasks that must occur simultaneously does exist, but the basic constructs of the language makes its use difficult. I wrote nearly all of the agent in Python, which, in recent versions, offers the use of *generators*, a construct that is close to the older technique of *coroutines*. Coroutines are a generalization of subroutines that enable exit and entry of a function at any point, not just the top and bottom as in subroutines. Python, at least in the version I used (2.4) does not offer full support for coroutines, as you cannot pass values back to the function. However, the ability to return to the "middle" of a function on the next call is used extensively to provide a lightweight form of task switching without the overhead of thread creation.

In a world with a programming language that would be ideal for this task, there would be no difference between the part of the agent that listens, the part of the agent that decides, and the part of the agent that does. Memory would be distributed across the entire agent with minimal access restrictions. The diagram would be less an image of boxes connected by arrows, and more a semi-differentiated goo that could expand, contract, and mutate at will. Instead of containing the world within the boxes, the agent system would be able to push against the porous, flexible boundary of the goo, growing (in abilities, if not size) as needed.

Yet such a language does not exist, nor do I know how to massage current languages into the form I outlined. Thus at some level I have to decompose my agent system into a manageable series of parts. As I will detail in the sections below, these parts are interconnected in numerous interesting ways, converting the system diagram into at least boxes with many recurrent arrows, if not a diagram approaching goo.

THE PARTS OF THE AGENT

The agent for *syngvab* is built from four components:

- *Listen*: takes in audio data, calculates necessary features, determines whether a sound is speech or non-speech

- *Consider*: takes in non-speech audio data and runs it through a series of neural networks, resulting in commands to control the creature

- *Choose*: takes the outputs of the neural networks and decides which one to use

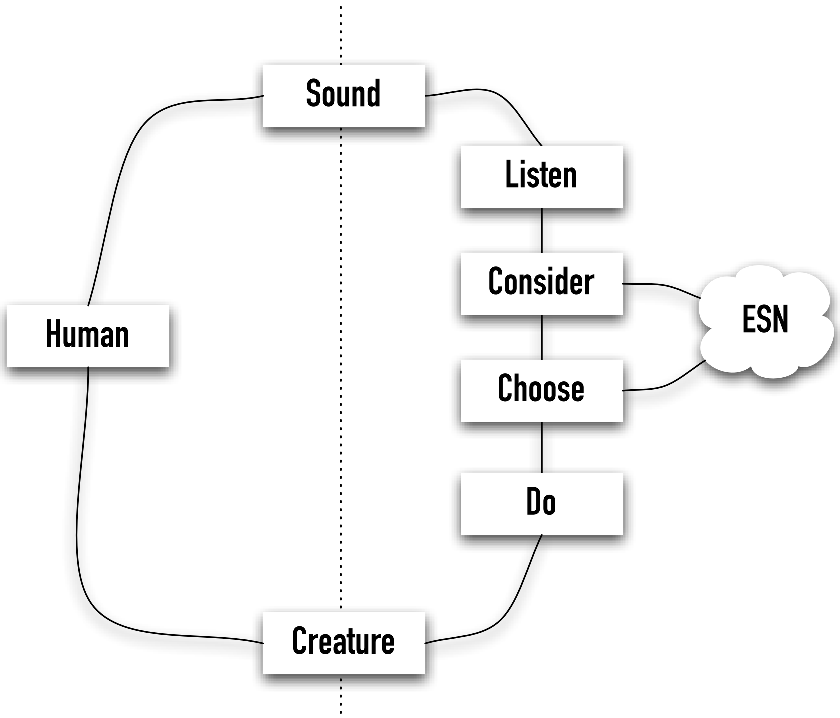

Figure 7.1: Schematic of the agent design. To the right of the dotted line represents parts of the agent, to the left represents parts of the human.

- *Do*: takes the commands and translates them into outputs for the simulated and physical creatures

Figure 7.1 is a simplified diagram of the agent decomposition.

The following sections describe in detail the workings of each subsystem of the agent. Keep in mind the goal to smoothly connect each subsystem with the next. While I present each system in a linear fashion (and indeed, each is called sequentially), the desire is otherwise and I am always looking for a new way to represent the systems of the agent to better match this want.

7.2 LISTEN

The "front" of the creature for the *syngvab* agent, as it might be called, is the *Listen* system. The agent needs some way to take in sounds from the outside world, and specifically from people interacting with it, and convert those sounds into meaningful values for the later parts of the system, specifically *Consider* . Of course this necessitates a series of *decisions* from my point of view as to what features of the incoming

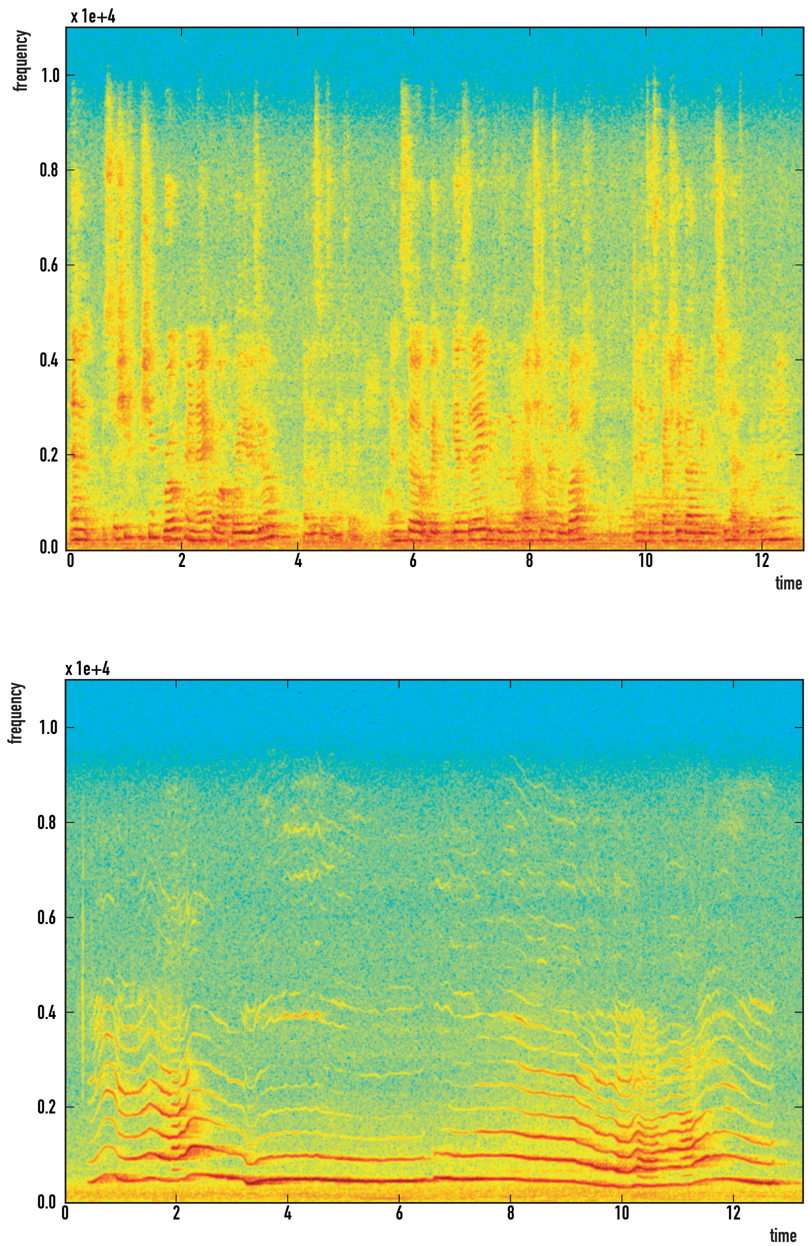

Figure 7.2: Spectrograms of speech and non-speech signals. *Top*: speech. *Bottom*: non-speech. Notice the slower changes in the non-speech signal, along with the smaller number of harmonics.

audio stream are important to consider; these choices will dictate and limit the later possibilities of the creature and the agent, thus making the selection of measurements an important one.

Indeed, the number of options for potentially useful measurements of audio signals is virtually unlimited. In the proceedings of signal processing conferences one can find new features with ever-more-complicated names that purport to offer solutions to vexing problems of audio analysis. This is complicated by the dual representation of audio data: on the one hand, the time domain, and on the other, the frequency domain. In practicality, there is a severe trade-off inherent in our choice of time and frequency resolution. Since we move to the frequency domain through the short-time Fourier transform (STFT) we can have high temporal resolution and low frequency resolution, or vice versa, but not both. Yet this is not as much of a problem as it seems; if we take a reasonable window size we will not have problems capturing the necessary and important frequency transients in the signal of interest.

We can see a graph of our two types of signals in Figure 7.2. The top part of the figure is one example of a speech signal; the bottom, a non-speech signal. Globally we can see how the speech-signal seems to have more transients in the temporal domain; that the non-speech signals seems to change over a slower time window; and the speech signal has a larger number of high-frequency components. How, then, do we represent this information in a way that we can manipulate in our agent?

PERCEPTUAL FEATURES FOR AUDIO ANALYSIS

Even with the large proliferation of possible audio features to consider, there are a set of features that are certainly more common than others (Peeters 2004), and this thesis will draw from this limited range, with the understanding that with more time and resources consideration of other esoteric options might be better for my particular application.

For reasons of computational expediency, I work entirely in the frequency domain, deciding not to calculate any time domain features; thus, once we have computed the STFT of the signal we can focus immediately on the features that are derived from that signal. Depending on our window size the number of possible frequency bins that come as a result of the STFT can be quite large. Additionally, the human perceptual system is not equally sensitive to all frequency bands, meaning that we can filter and warp the results of the STFT to better match the features that are most important for the workings of our cochlea and auditory cortex. We do this through the choice of *perceptual* audio features. These features and calculations remove information that is not used by the perceptual system in order to better match in a *machine* context what we do naturally in a *physical* context. Perhaps the most well-known perceptual audio algorithm is the well-known MP3 compression; this algorithm works by removing

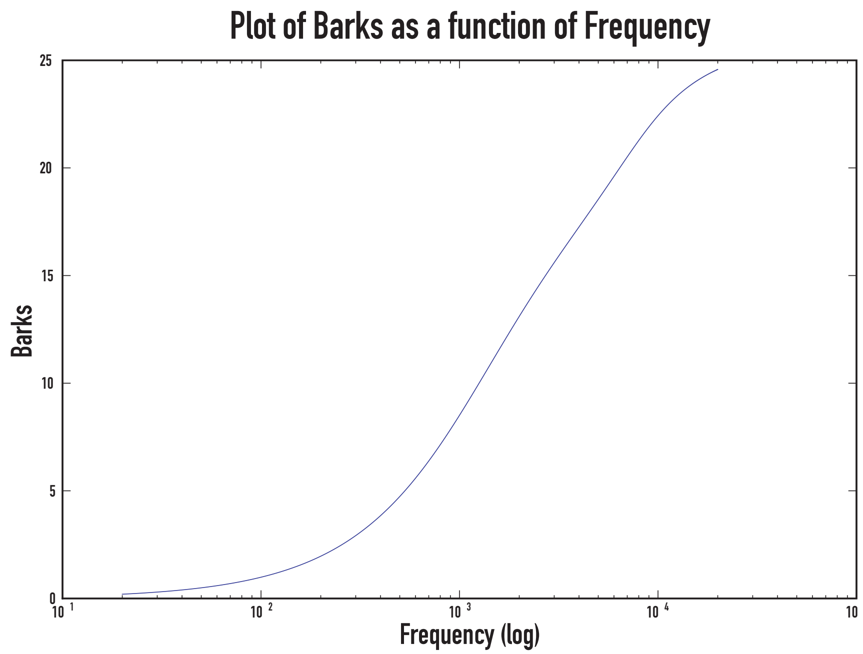

Figure 7.3: Graph of Barks as a function of frequency. The x-axis is on a logarithmic scale.

frequencies above the range of hearing, as well as those that are masked either in the temporal or frequency domain (Painter and Spanias 2000). Recently a number of researchers have shown how the use of perceptual features can improve clustering and segmentation of audio data, as well as be used in re-synthesis and warping across performers, styles, and genres (Ellis, Whitman, Berenzweig, and Lawrence 2002; Berenzweig, Logan, Ellis, and Whitman 2004; Jehan 2005; Whitman 2005).

The first step in computing perceptual features from our audio is to warp the raw STFT in order to reflect the frequency response of our cochlea. We can warp the STFT using a variety of filters, the most common coming from the Mel scale and giving us a set of what are known as Mel-frequency cepstral coefficients (MFCC). However, Zwicker and Terhardt (1980) showed that a better set of coefficients to use are the *Bark bands*; they can be approximated by

$$B = 13\arctan\frac{f}{1315.8} + 3.5\arctan\frac{f}{7518} \tag{7.1}$$

with f in Hertz and with a graph of frequency versus Barks found in Figure 7.3

This representation of the data in the frequency domain is what we use in all of our later calculations. I chose four features to use for our analysis: total loudness, spectral centroid, sharpness, and spectral flatness (Peeters 2004). These represent many of the most commonly-used in perceptually-based analysis.

SUPPORT VECTOR MACHINES

The Support Vector Machines (SVM) classifier is a widely used method for creating a decision boundary between two classes of data (Hastie, Tibshirani, and Friedman 2001, pp. 371-389). The procedure works by finding the *hyperplane* that best separates the data in a high dimensional space (Figure 7.4, top). In the perfectly linearly separable case, this is relatively easy. However, there are often times when it is not possible to find the hyperplane that separates the data. In these cases one can implement *slack variables* that control how much one tolerates data from the other class appearing on the wrong side of the boundary (Figure 7.4, middle). This helps to create boundaries that work well in practice, even if the data cannot be completely separated by a hyperplane.

In common with other types of machine learning techniques, we can get around the problem of linear separability by mapping the training and testing data (of dimensionality N) into a higher dimensional space of dimension P (where $P > M$) and evaluating the SVM algorithm there. A perfectly separable linear hyperplane in the space of dimension P will map to a non-linear decision boundary in the space of dimension N. This enables one to find good separation between classes even when the data cannot be easily divided in the native space (Figure 7.4, bottom).

I use a widely-used SVM library called `libsvm` (Chang and Lin 2007). The library and its associated utilities allow easy scaling, training, and prediction of SVM results[1]. Scaling of the parameters to fit within the bounded range $[-1, 1]$ is necessary to prevent potential numerical difficulties, as well as ensuring that larger numeric ranges do not dominate smaller ones (Chang and Lin 2007, p. 4). I stayed with the default choice of the radial-basis function (RBF) kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(\gamma \parallel \mathbf{x}_i - \mathbf{x}_j \parallel^2); \tag{7.2}$$

this, however, necessitates choices for the optimal parameters of γ and C, the generalization "knob".

The training and testing data came from the recordings I made of colleagues making "non-speech" and "speech" sounds that I described in Chapter 2. There were five different speakers, 3 female and 2 male. There were approximately 10 examples

[1] Note for those of you trying this at home: I found that the model file produced by the command-line utilities was not able to be used properly by the python bindings. The support vectors were exactly the same in the two cases, but there were different labels. This meant that I had to do training and prediction all using the same bindings, rather than using a model file made from the command-line utilities in the python scripts.

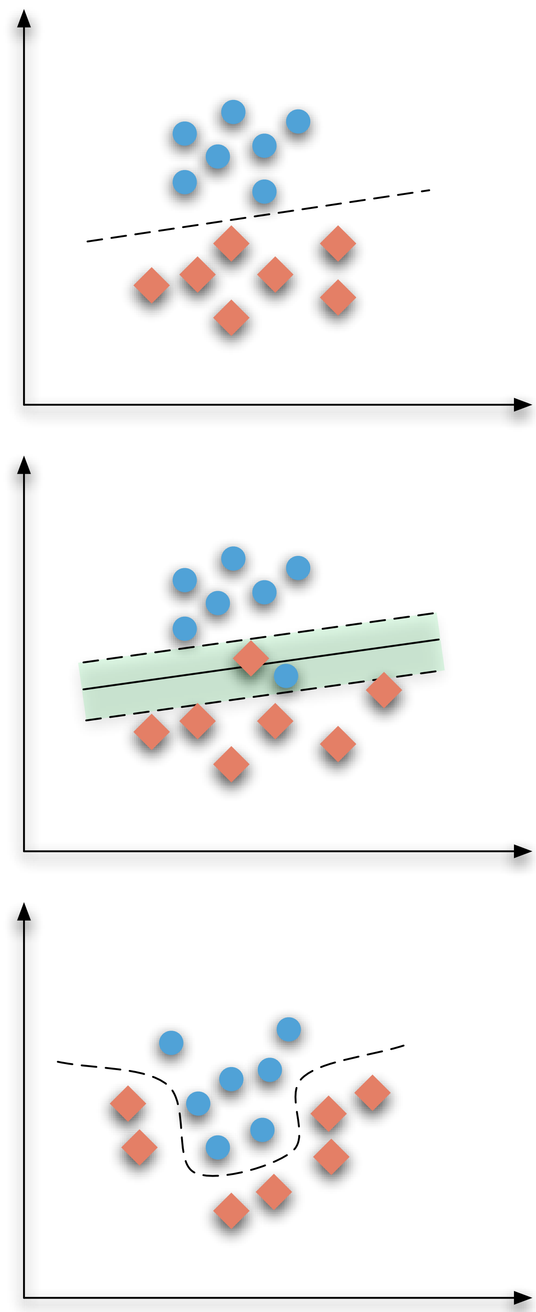

Figure 7.4: Cartoon representation of separation of classes using SVMs. *Top*: Perfectly linearally separable case. *Middle*: Non-perfectly separable case using "slack" variables and a margin. *Bottom*: Non-perfectly linearlly separable case using kernel functions of a higher dimension.

of speech and non-speech for each speaker, comprising a total sum of around five minutes of total audio data. I split the sum of data into training and testing sets: approximately 40% for training and 60% for testing. I used the included `grid.py` python script to do a grid search of parameters for the radial-basis function kernel using five-fold cross-validation. This grid search process was repeated three times, each time narrowing the search space based on the found parameters. I ended up with setting $C = 2^{19}, \gamma = 2^{3}$.

After this procedure, I had what would be considered quite amazing results: in the training phase, my prediction accuracy with the best parameters was 91%. When using this trained model on the testing data, the accuracy dropped to only 89%. Considering that I collected the data in a variety of environments with only the microphone available on my laptop, I am extremely pleased with these prediction values. One might think that better results could perhaps be found with more controlled recording environments and better choice of microphone. However, this would ultimately lead to degenerated performance *in-situ*, as I cannot predict nor control the acoustic environments of participants' homes. Thus, the best choice would be to record in a variety of places using off-the-shelf, consumer microphones to better cover the wide range of possible acoustic locations.

Once we have determined whether or not a given frame of sound is speech or non-speech, we pass on the four sets of computed perceptual features to the next system, the *Consider* system. From here, we have to consider how to use the values to cause *syngvab* to move.

7.3 CONSIDER

By this point we have determined, through the *Listen* system, whether or not a given sound (frame) is non-speech. Via this, we also have a set of values for interesting and relevant perceptual audio features. Now that we have that data, what do we do with it? We need to create some way of using that sound and turning it into motion. This is where one would normally speak of the *mapping* problem: how to take one arbitrary piece of data and turn it into another arbitrary piece of data. Put mathematically, this is the relationship of $\mathtt{output} = \mathtt{f}(\mathtt{input})$. Yet this functional relationship, while perhaps true *programmatically*, in fact represents the opposite of how we want to consider the situation *conceptually*. This relates to our discussion of technological development in Chapter 4: our possibilities are bounded[2] by the way the problem is posed. If we think of our problem here as a simple "mapping" of one set of data to another, we will perhaps ignore other options that connect the disparate sets of data in more interesting ways.

[2] This boundary is indeed soft and (sometimes) permeable, but represents generally accepted limits.

The computer music community has oft dealt with this issue. Until quite recently the aural output of a musical instrument was directly and intimately related to its means of production: air blown into a tube, a string excited by a bow or finger, a drum struck by a hand or mallet. The development of the MIDI interface in the mid 1980s, however, decoupled the input and output: now striking a key on a piano could just as easily make the sound of a flute as the sound of a piano. Interactions with any object could give rise to any sound as long as a common language was spoken. The function f could be anything, allowing one-to-one, one-to-many, many-to-one, and many-to-many connections between input and output data. Because of the abstracting nature of digital technology, anything can become anything else. This is an impossibly large space of possibilities, so expansive as to require the imposition of some limits. However, as I will soon note, the limits that have been placed restrict our choices to an unbearable extent.

In his recent thesis Marc Downie developed arguments against a functional concept of mapping, suggesting that there is a "need for a greater range of vocabulary, for greater nuance, for describing this very struggle and organizing the intellectual field around it" (Downie 2005, p. 30). Indeed, all too often discussions about how to connect some measurement to an output describe a simple linear relationship: as I do more of something as an input, I get more of something as an output. And vice versa. Such a sadly limited use of the relatively powerful possibilities of digital technology severely limits one as the designer of these systems. I would argue, like Downie, that this result comes from the purely functional view of this problem. If we let some type of autonomy come into play, if we allow a bit of *unpredictability*, then we at least perceptually and conceptually break the function apart. Input and output can mix together in ways that are unknown to us as the designer.

This is all the more important as we consider movements for *syngvab*. I have no real *a prior* expectations of how the creature will move given a particular non-speech sound. Moreso, I have no real predictions as to what sounds people might make, given my strange request and the situation I am creating for them. How do I represent what I cannot know or cannot anticipate? How can I make the connection between a given sound and a given output without knowing what the input might be? Unlike in some of Downie's earlier work, as well as the projects of the Synthetic Characters group at the MIT Media lab (Downie 2000; Blumberg, Downie, Ivanov, Berlin, Johnson, and Tomlinson 2002), I cannot create representations for *syngvab*. There is no "sit", no "follow-the-leader", no abstract concepts that we might try and connect to a percept in a many-to-one fashion. This has analogies with the TGarden research project of Sha Xin Wei and colleagues. In their work with gesture they eschew any attempt at *recognition*, and rather focus on *tracking*; less the creation of an abstract model (as in Downie's work) and more the direct use of the action (gesture) for meaning (Wei 2002, pp. 455–457).

This desire puts us in a difficult conundrum, similar in magnitude to what I faced with the design of *syngvab* as a physical object. If we could explore the space of possibilities, however, in some sort of interactive fashion, we might be able to constrain our options in a data-driven manner. To take a concrete example: if I have some samples of non-speech sounds and can observe the outputs of the system for a number of different trials, then I might be able to influence the parameters of the system so that when a new example is heard, the system produces an interesting output. This is the approach in the agent here. The process is split between the *Consider* and *Choose* systems. In the *Consider* system, I train a type of Recurrent Neural Network (RNN) through an interactive evolution process that occurs in the *Choose* system. This coupling of the RNN with an evolutionary search for weight matrices enables me to quickly search an impossibly vast surface of options. In the end, the trained RNN is indeed a black box: I do know know how the inputs are translated into outputs. I know what the limits of the network are and how those limits bound my options. But the details of mechanism are nonetheless unimportant: what is meaningful is the perceptual outcome, the result.

We can know jump directly into the messy and abstract details of the implementation. What I hope will be clear is how my choice of the particular RNN is related to these concerns of mapping. Once we come to the discussion of interactive evolution, in Section 7.4, we will see how these two systems link together to help put down any purely functional mapping.

ECHO STATE NETWORKS

Neural networks are common tool for various types of computer pattern recognition, generalization, and prediction tasks. The most common type of neural network is the Feed-forward Neural Network (FFNN), a diagram of which can be found in the left of Figure 7.5. In traditional FFNNs information "flows" in only one direction: data is presented to the input units, which are in turn connected to units in a hidden layer, which are themselves connected to units in the output layer. Neural networks can be considered as graphs, with the units of the network as nodes and connections (weights) between units as arcs. Thus, in a FFNN, there are no cyclic paths between any nodes. Training of FFNN often is through gradient descent or back-propagation. In their standard form FFNNs are not able to represent temporal information; however, this can be overcome by using a multiple of the input units to present to the network temporally delayed information.

A RNN is a generalization of feedforward networks, allowing recurrent connections from one node in a higher layer to one in a lower layer (right of Figure 7.5). Weights between nodes are allowed in both the forward and backward directions; additionally, self-connections are possible. Such a topology enables forms of temporal

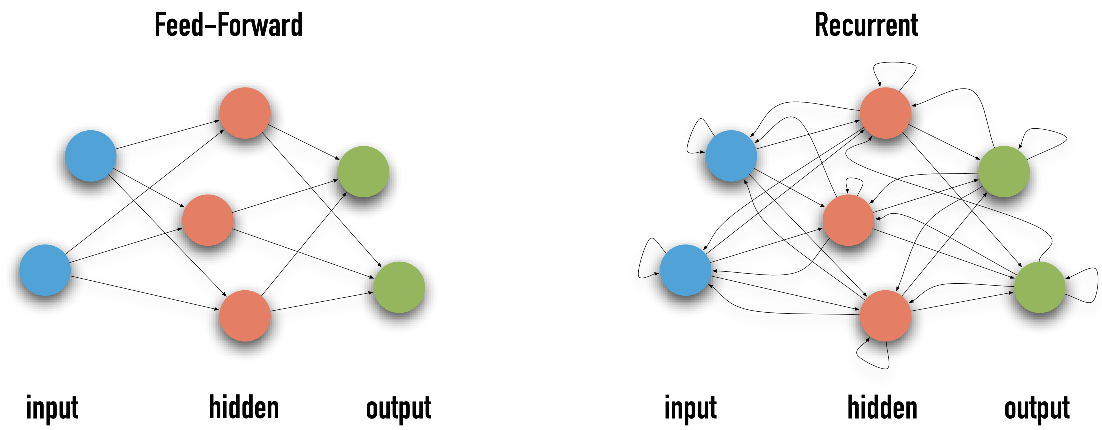

Figure 7.5: Common neural network topologies. *Top*: feed-forward. *Bottom*: recurrent. In the feed-forward case information "flows" from the input layer through the hidden layer and to the output layer, without any "backwards" travel. In the recurrent case, information can flow forwards or backwards; as well, self-connections can exist. Because of their structure, recurrent neural networks are able to represent temporal information in a way standard feed-forward networks cannot.

memory by way of the recurrent links; weights that connect backward encode how one node relates to another in a temporal sequence.

The practical problems with the RNN are many, however, making its use less a panacea and more of a frustration. Briefly put, they are:

- What is the right structure of the network? How sparsely do we make the connections?
- How fast does the network converge?
- How do we prevent overfitting?

While there have been reasonably useful solutions to each of those problems alone, each has severe drawbacks. Jordan and Ellman networks enable us to address the first concern, but the possible answers to the second leave us less satisfied, with the common training algorithms being extremely computationally taxing.

Recently, however, there as been the development of an alternative formulation of the RNN the echo-state network (ESN) (Jaeger 2001*b*; Jaeger 2001*a*). ESNs enable us to focus only on the training of the output weight matrices through what reduces to simple linear regression, solving the problems of network topology determination and computational cost of network training. ESNs have recently been used in a number of situations for difficult non-linear problems, including digital audio modelling

(Squartini, Cecchi, Rossini, and Piazza 2007) and, most relevant for this thesis, robotic motor control (Ishii, Zant, Becanovic, and Ploger 2004; Plöger, Arghir, Günther, and Hosseiny 2004; Salmen and Ploger 2005). Additionally, and most relevant for our discussion about the *Choose* system later, is recent work that used evolutionary techniques to determine network parameters (Ishii, Zant, Becanovic, and Ploger 2004).

In a general recurrent neural network the distinction between input, hidden, and output layers can become quite muddy; however, these names are meaningful when we write of the network, so we continue to use them. Consider a network[3] that has the standard three sets (or layers) of nodes: K input nodes $\mathbf{u}$; N internal (hidden) nodes $\mathbf{x}$; and L output nodes $\mathbf{y}$. At any given discrete time t, the vectors can be denoted as $\mathbf{u}(t) = (u_1(t), \ldots, u_K(t))$, $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))$, and $\mathbf{y}(t) = (y_1(t), \ldots, y_L(t))$. We can represent the connections between nodes by a series of weight matrices with components $w_{ij} \in \mathbb{R}^1$, where $w_{ij} = 0$ indicates the absence of a connection between nodes i and j, and $w_{ij} \neq 0$ indicates the strength of the connection between nodes $j \rightarrow i$. We can then define the weight matrices as follows: $\boldsymbol{W}^{in}_{N \times K} = (w^{in}_{ij})$ for the input weights; $\boldsymbol{W}_{N \times N} = (w_{ij})$ for the hidden layer weights; $\boldsymbol{W}^{out}_{L \times (K+N+L)} = (w^{out}_{ij})$ for the output weights; and finally, $\boldsymbol{W}^{back}_{N \times L} = (w^{back}_{ij})$ for the back projection weights from the output layer to the hidden layer. There are no restrictions on the location of the weights in the matrices, thus allowing any level of recurrency and direct connections from input to output and from one output node to another.

We move this network one step forward in time by two equations. First, we calculate the hidden node activation vector:

$$\mathbf{x}(n+1) = \mathbf{f}(\boldsymbol{W}^{in}\mathbf{u}(n+1) + \boldsymbol{W}\mathbf{x}(n+1) + \boldsymbol{W}^{back}\mathbf{y}(n)) \qquad (7.3)$$

The set of functions $\mathbf{f}$ are the activation functions for the hidden units in the network, commonly set uniformly to $\tanh$. Calculation of equation 7.3 is called *evaluation*. Next, we update the output nodes:

$$\mathbf{y}(n+1) = \mathbf{f}^{out}(\boldsymbol{W}^{out}\,[\mathbf{u}(n+1), \mathbf{x}(n+1), \mathbf{y}(n)]) \qquad (7.4)$$

Similar to Equation 7.3 the functions $\mathbf{f}^{out}$ are the activation functions for the output nodes. The notation $[\cdots]$ denotes horizontal concatenation of the $\mathbf{u}$, $\mathbf{x}$, and $\mathbf{y}$ vectors. Evaluation of Equation 7.4 is called *exploitation* of the network.

The echo state property, which I will describe below, allows the $\boldsymbol{W}^{in}$ and $\boldsymbol{W}^{back}$ matrices to be set with few restrictions; these matrices are commonly full and randomly initialized from a uniform $[-1, 1]$ distribution. For the hidden weight matrix

[3] A network can also be considered a graph, with weights being the arcs that connect nodes together.

$\boldsymbol{W}$ there are minimal conditions to ensure that the network obeys the echo state property, as well as does not veer off into chaotic territory:

1. Randomly initialize the hidden weight matrix $\boldsymbol{W}_0$ from a uniform distribution ($w_{ij} \in [-1, 1]$); the matrix should be sparse, with density around $5 - 10\%$.

2. Calculate the eigenvalues of $\boldsymbol{W}_0$ and find the one with maximum absolute value λ_{max}. Normalize the weight matrix using this value: $\boldsymbol{W}_1 = (1/\lambda_{max})\ \boldsymbol{W}_0$.

3. Scale this new matrix by a desired spectral density α, where $\alpha < 1$: $\boldsymbol{W} = \alpha \boldsymbol{W}_1$.

We can now say that the network obeys the echo state property. A proof of this is beyond the scope of the thesis, but can be found in Jaeger (2001*b*).

Given that we now have an echo state network, we can see that the network is independent of the specific values of $\boldsymbol{W}^{in}$ and $\boldsymbol{W}^{back}$, allowing these weight matrices to be set according to the intuition of the network designer. Thus, in addition to tuning the structure of these matrices, the only other free parameters are the dimensionality N of the hidden layer, and the choice of spectral density α. We can see that α sets the time scale of the network: small alpha, and the network responds quickly to changes in input, large alpha and inputs tend to be damped on longer time scales.

All of this preparation leads us to the surprising conclusion that to train the echo state network all we have to do is train the output weight matrix $\boldsymbol{W}^{out}$. This training procedure is simple and boils down to linear regression:

1. Generate the training data and store it in input and output vectors $\boldsymbol{u}_{train}(t)$ and $\boldsymbol{y}_{train}(t)$, respectively.

2. Drive the network with each $\boldsymbol{u}_{train}(t)$ and $\boldsymbol{y}_{train}(t)$, storing the results of the hidden layer vector $\boldsymbol{x}$ in a collection matrix $\mathbf{M}$.

3. Collect also the inverted training outputs: $\mathbf{C} = \tanh^{-1} \boldsymbol{y}_{train}(t)$.

4. Compute the pseudo-inverse of $\mathbf{M}$ and solve for the trained output weight matrix: $\boldsymbol{W}^{out} = (\mathbf{M}^{-1}\mathbf{C})^T$, where $(.)^T$ denotes the matrix transpose.

We now have a trained network. We can use the weight matrices, along with Equations 7.3 and 7.4 to calculate the outputs given any particular input vector.

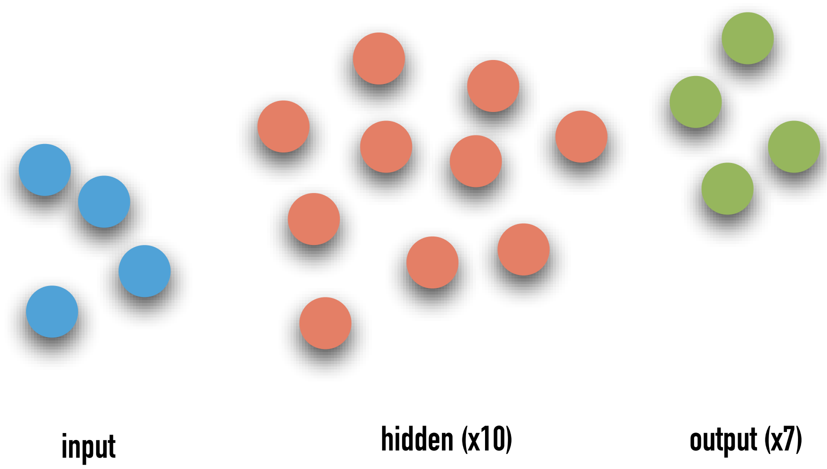

Figure 7.6: Schematic of the ESN for *syngvab*

APPLICATION OF THE ESN TO *SYNGVAB*

The ESN topology for *syngvab* is relatively straightforward (Figure 7.6. The input layer consists of four nodes, one for each feature that came from the *Listen* system. The hidden layer consists of 100 nodes; this number was kept small for computational efficiency. Its density is only 0.1. The output layer consists of 28 nodes, four for each motor: steps, acceleration, deacceleration and speed.

I use the ESN in two different ways for determining the control of *syngvab*. In the *offline* mode, I directly couple the network to the evolutionary development of the output weight matrix that is implemented in the *Choose* system. I present a particular sound to the inputs of the network and cycle through a set of possible networks, each with different topology and output weight matrix. The *Consider* system determines the outputs that are then passed to the *Do* system for implementation on the simulated or physical creature. These outputs are then rated and given a fitness value, enabling the *Choose* system to move forward one step. In the *online* mode the inputs move through the network producing an output that is then directly sent to the *Do* system. The only use of the *Choose* system is to select which of the trained networks is the currently active one.

In both modes I scale the network inputs to lie between -1 and 1. The outputs of the network, lying in the same range, are then scaled to physically meaningful parameters.

7.4 CHOOSE

To create movements for *syngvab* in the *Choose* system, we use draw from the techniques of genetic and evolutionary algorithms. These techniques, in turn, draw from a biological metaphor of development. By framing our problem in a particular way, we can use the random search of mutation, crossover, recombination, and (a-)sexual selection to create new "organisms" that better match a given *fitness function*. This function encapsulates the desired goals of the particular genetic algorithm, and our selection process uses the fitness function to remove those organisms from the population that fare poorly, choosing instead to breed the organisms that perform the best, given the fitness function.

Evolutionary development of control for robotics has a long history (Nolfi and Floreano 2000) and can be roughly divided into two (partially) overlapping segments: the evolution of control for robots (Hurst and Bull 2006; Gruau and Quatramaran 1996; Watson, Ficiei, and Pollack 1999; Lipson 2006; Yanase and Iba 2006; Lund, Miglino, Pagliarini, Billard, and Ijspeert 1998; Baldassarre, Nolfi, and Parisi 2003; Lewis, Fagg, and Solidum 1992; Jakobi 1998); and the evolution of robot morphologies (Sims 1994; Lipson and Pollack 2000; Pollack, Hornby, Lipson, and Funes 2003; Macinnes and Di Paolo 2004; Lipson 2006; Dittrich, Burgel, and Banzhaf 1998; Lund 2003; Bongard, Zykov, and Lipson 2006). While it might have been interesting to evolve the morphology of *syngvab*, in this thesis we focus on the evolution of a control system.

What happens, however, if in our problem we cannot design an appropriate fitness function? We can turn to interactive evolution, the use of a subjective human response as the fitness function. Interactive evolution and genetic design have both been used extensively for the development of *aesthetic* projects. Karl Sims was one of the first to do this in the area of computer graphics, showing how new types of scenes and images could be created through the interaction of human and algorithm (Sims 1991). Since then there have been a large number of other projects that use the human as the fitness function for some pre-determined aesthetic goal (Kim and Cho 2000; Lapointe 2005; Dahlstedt 2001; Chen and Miikkulainen 2001; McCormack 2001; McCormack 2002; Johnson 1999; Takagi 2001; Biles 1999; McCormack 2006; Lapointe and Époque 2005; Draves 2005; McDermott, Griffith, and O'Neill 2005). Using humans as the fitness function for these types of evolutionary goals is the only principled way to use genetic techniques for many types of design, as there is no single means of measuring aesthetics in a quantitative way. Better to use the subjectivity of a human for development, rather than an abstracting mathematical structure that fails to capture anything interesting about aesthetics.

While the movement of *syngvab* is not singularly aesthetic in nature, there is a quality of it that demands such a stance. As I described in the previous section, connecting non-speech sound to movement of a non-zoomorphic creature demands some type of black-box approach. Thus, in determining the types of movement, we can only be guided by desire and spontaneity, hoping to steer development in interesting ways but using the power of randomness to suggest possibilities hitherto unknown.

APPLICATION OF GENETIC ALGORITHMS TO *SYNGVAB*

The use of genetic algorithms in the development of a movement profile for *syngvab* is relatively straightforward. As mentioned in the previous section, the "organism" for our algorithm consists of the raw ESN output matrix. Due to known problems of neural network topology in evolutionary algorithms, we do not use the crossover operation (Angeline, Saunders, and Pollack 1994). The population of weight matrices consists of four organisms, the small size needed for expediency. At each generation step, I take the current output weight matrix for each organism and display its result in the *syngvab* simulator using one of the sounds collected earlier. I rate the "interestingness" and "appropriateness" of movement, which is then used in the subsequent fitness calculation. The top creatures are selected by the algorithm and a small subset (10%) of weights are mutated by a small amount ($\pm$ 0.1), drawn uniformly. For the control system in this thesis I continued this process for 30 generations[4]. All of these generation steps take place offline.

The online part of the *Choose* system is simple; it merely selects one of the four evolved networks in a random fashion, thus ensuring that there is some semi-arbitrary response and forcing people to continually rethink their interactions with *syngvab*.

7.5 DO

After the previous three systems of the agent, we come to what is, in some ways, a disappointing denouement, for the *Do* system is extremely straightforward. Because of the design choices we made in the construction of the physical creature and its simulation, as well as how we scale the outputs of the ESN, the translation of the outputs of the *Choose* system is quite easy.

The *Do* system consists of an output manager that collects commands to send to the different types of creatures, and a set of output classes that all implement a common interface. One class takes commands destined for the simulated creature, the other output class translates commands for the physical creature. In this way the output manager can be ignorant of the technical details of sending commands

[4] In a more developed methodology, we could perform this interactive process with a set of naïve observers, thereby collecting and considering a large space of "interesting" options.

to the various outputs, and these details can be encapsulated and kept near to the appropriate output. The output manager implements a command queue that, when the `process()` method is called, sends the command to each of the associated outputs in the manager. Each output in turn translates the given command into the form necessary for the output.

SIMULATED CREATURE OUTPUT

For the simulated creature there is no translation of outputs. I designed the command system for the simulation to take the same form as the physical creature, so all commands sent to the simulated creature go through their own internal translation to move from steps and accelerations/deaccelerations to durations and forces. I communicate with the simulated creature over OSC.

PHYSICAL CREATURE OUTPUT

The physical creature implements all of the commands described in the Table 6.5 in the previous chapter. Thus, we merely have to take a textual command such as SET_MOTOR_SEQUENCE_PROFILE and a series of integers and convert it to the appropriate byte stream. This is done through simple bit shifting, taking a number such as 500 and turning it into a two-byte value.

The motor control boards work only on unsigned values, so we have to determine which step values might code for reverse directions and send the correct SET_MOTOR_SEQUENCE_DIRECTION command prior to sending the profile values.

The firmware on the creature turns the profile command into the sequence of accelerations, speeds, and deaccelerations necessary for moving the creature as desired.

⁂ ⁂ ⁂

Now we have described both the physical design of the creature, and the design of the software agent. It is time to look at what happens when we put *syngvab in-situ*, how people react to it on a semi-long-term basis.

8

Study of the Robotic Marionette in the Home

Now we have come to the study of *syngvab*, one of the main goals of this thesis. This is where traditional AI would spend its most time, with the belief that the only important part of a theory is its actualization: "if your alternative is so good then you will use it to write programs that solve problems better than anybody else's, and then everybody will believe you" (Agre 1997, p. 149). While I agree in part that building things and studying them in the wild is necessary, and indeed wrote of it at length in Chapter 4, I am not willing to simply cast off the entirety of the earlier parts of this thesis if the use of *syngvab* is not as expected. The theory that I drew from for the thesis is independent in its own right, while also serving to support the creation of *syngvab*. In my work theory and practice mutually support and reinforce each other; each can exist own its own, but the disparate mingling, the messy combination of both strengthens the research enterprise as whole.

One of the biggest challenges of this thesis was the *in-situ* study of *syngvab*. In order to do this I had to make the system robust enough to enable regular, unattended use. As we will see, the project somewhat fails in this respect, illustrating the difficulty of taking a research project out of the laboratory. However, I would suggest that this is vital to the future of the Human-Computer Interaction (HCI) enterprise, especially as we begin to develop more and more of these relational artifacts. Our creatures will remain impotent and unaware of the difficulties, the messiness of real-life unless we force them to interact with people in natural settings.

8.1 PROCEDURE FOR STUDYING *SYNGVAB* IN THE HOME

I gave *syngvab* to a small set of participants for use in their home for around three days each. I asked each person to partake in three parts of the study. The first part was a pre-interview session with myself; I explored the participant's knowledge of robotics, their understanding of their own production of non-speech sounds, and their desire for interaction with a robotic creature. I then explained, on a general level, the workings of *syngvab*, showing the participant the switches needed to turn the creature on and off, and as well as one way of interacting with the creature. I then asked them to interact with *syngvab* for at least five minutes a day. After each interaction with *syngvab* I asked them to write briefly about their experiences in a diary-like setting; what was it like, how were things changing over time, what types of things they did with the robot. This was the second part of the study. The third part was a post-interview session; I explored their thoughts about their time with *syngvab*, what new things they learned about themselves, and their desire for continued interaction with *syngvab* or things like it. This interview was sometimes conducted via e-mail.

I focus on one case study in this thesis. As I will describe in Chapter 9, there is certainly more work to be done, more participants to interact with *syngvab*, but the explication of a single user here will suggest the power and possibilities that now exist.

8.2 NON-SPEECH AND ROBOTICS

Cara is a 27-year-old female living in the Boston area. She is a musician, performing in both traditional and contemporary ensembles. Unique to her is her own personal knowledge of the production of non-speech sounds:

> Q: You were talking earlier about the sounds you make yourself. Can you talk a little more about that, sort of describe what they are, if you can, or maybe even make them?
>
> A: Can I just make the sounds?
>
> Q: And also whether there is any sort of correlation between that sound and a particular type of emotional experience or feeling that you have when you make those sounds, and whether it's by yourself or around other people?

> A: Usually when I make sounds, it's by myself to myself, or when I think I'm by myself to myself. Like, if I'm sitting at my desk at work, and I'm concentrating on something, I'll talk to myself as well as make little sounds. Like I do sound effects for my computer. [Makes sounds.] Usually I'm happy and making myself entertained. Then, if I'm looking at something, and I make some kind of realization or discovery or whatever, I'll go [makes sound], and it's like oh, cool, I did not know that. I know when I'm frustrated, and I can't get something right, I'll go like [makes sound], instead of swearing or whatever. And often I go like [makes sound], and that's just like, happy, bouncing along, doing my thing. Most of the noises I make are happy noises, and a lot of the time I'm not really conscious of doing them, and other people tell me that I am doing them, so I don't really know why that is or why I make the noises that I do.

Cara sees how the use of these sounds, their conscious and unconscious production, reflects the potential power of the non-speech for expression of internal experience, of communication to oneself certain emotions and aspects of one's life that cannot be put into words:

> Q: Why do you think that, in your own impression or point of view, why do you think using non-speech sounds might enable someone to be closer to, or experience some emotional aspect of their lives in away that they would be able to with speech?
>
> A: Because there are some things that there just aren't words for. Or, maybe somebody doesn't have the vocabulary to express them. And they can discover that they can express those things by making sounds instead of finding words. If you don't have the words for something, sometimes it's too hard to actually explain it in words that you do know. Sometimes its easier to make noises, or gestures.
>
> Q: Do you think that would be something for communicating to people? Or communicating to themselves? Or not even communication?
>
> A: I don't see it as being used as communication. I think it would be more useful for people communication with themselves. Because, if you have these feelings that you don't know how to express in words, and if you have this object sort of stimulating these emotions and sounds out of you, then you might be able to do that. Yeah, I don't think communication. I think it would be too hard, because other people don't really understand it as much. I think it would be too hard to use it for communicating to people.

The sounds that Cara makes are directly related to the feelings that accompany them; they are linked in a way that is difficult, if not impossible, to split. This reflects a deep somato-sensory link between sound production and emotional experience:

> Q: When you just made those sounds, were you feeling anything similar to what you feel when you make those sounds when I don't ask you to do them, when they just naturally happen.
>
> A: When I make the happy noises, I feel happy, and when I make the frustrated noise, I feel a little frustrated. Definitely, if I just make the noises, I feel the feelings a little bit that go along with them. The happy noises especially.
>
> Q: Can you make those sounds [the happy ones] with a pouty face?
>
> A: I'll try, but I don't think I can. [Tries to make sounds.] No, I can't! No, no, because I think I've always made those sounds in association with being happy or being content or whatever. Now, if I do them out of context, then I'm going to automatically feel happy no matter how I try.

8.3 INTERACTIONS WITH *SYNGVAB*

Following the interview I gave *syngvab* to Cara for her use. After her first few minutes interacting with *syngvab*, she wrote this:

> The first time that I made the creature move, I was actually kind of scared. I wasn't expecting it to move the way it did. After watching it move for a minute, though, I started to laugh because I thought that it was really neat that it would move to certain sounds that I made. I tried making higher-pitched noises to see how it would move, and then I tried making lower-pitched noises to see if it would move differently. I noticed that it moved differently to each noise that I made. I was actually laughing a lot while interacting with the creature. It made me feel really happy to see something responding to my sounds, and it was also amusing. The more it moved, the more inspired I became to make noises, and I also got more creative with the different noises that I made. The physical appearance of the creature did not incite any emotions in me at first, but after it began to respond to my sounds I immediately began to feel like it had a personality and character.

I needed to perform some maintenance on *syngvab* after these early interactions; because of present design and control problems, *syngvab* would easily get caught at the top or bottom of the stage, necessitating a manual placement of the creature. This unexpected interaction did not seem to cause problems of Cara and the creature, however, as later diary entries reflect:

> I started out making high pitched noises to the creature and it responded by promptly attaching itself to the ceiling of its box. This time with the creature I was not startled by its movement like I was the last time. I enjoyed watching its response to my voice, and I found it amusing. This time, I had preconceived ideas about how the robot would react to my voice, and I had even associated some personality traits with the creature. It reacted as I expected it would, and it reacted differently to each high or low pitched sound that I made.

We can see the beginning of a deep connection between person and object forming. Cara is not as "startled" by the movement of *syngvab* in these later interactions; she even begins to ascribe some "personality traits" to the creature. She is starting to form a relationship with *syngvab*, understanding how the creature will move given certain sounds.

8.4 AFTER THE INTERACTIONS WITH *SYNGVAB*

I conducted a post-session interview with Cara over e-mail. Her responses show the power of interactions with *syngvab*, reflecting the design choices made earlier. We first see how Cara was part of a creature-person dyad, with the movement of *syngvab* encouraging her to try different sounds:

> Q: Although you had only a short time to interact with syngvab, do you think the creature influenced you to make different sounds, or did you try to influence the creature's movement through your own sounds?
>
> A: -yes, he did. after a minute or so, i started experimenting with different sounds to see how syngvab would react. i experimented by making new sounds and i also tried to see how he would move in response to different pitches of sounds and different volumes.

Even so, we should take note of the negative aspects of experimental, the possibility of the movement and interaction being initially surprising or frightening:

> Q: Was there anything surprising about the interactions with *syngvab*? Not necessarily that it moved in a surprising way (although that's interesting too), but was there anything about how you reacted to it that

> was surprising for you? A: -although i had seen syngvab move before, i was surprised by my initial response to the creature when i made him move myself. i was a little frightened when syngvab first responded to the sounds that i made, maybe because i wasn't sure what it would really do.

This is an extremely important point to consider; when we attempt to create these deeply personal interactions between person and machine, the initial experiences, the beginning of relationship formation, are crucial to later understanding. Perhaps the design of robotic creatures should take into account a learning phase that attempts to reduce the potentially frightening aspects of the encounter, those parts of robotic movement that might be too surprising for someone unacquainted with robotic motion.

But after the frightening aspect of the interaction passed, we can see how there was the beginning of relationship formation:

> Q: Do you think you were forming a "relationship" with *syngvab*?
>
> A: yes - i'm not sure why or how i reached the conclusion that syngvab is a male, but i keep referring to it as "he." to me, as soon as he started moving, i began to imagine him with a personality, character, etc.

Even so, this does not mean that Cara desires for *syngvab* to be around her apartment constantly:

> Q: Do you think you will miss having *syngvab* around?
>
> A: i'm not sure i'd miss him because, like i said, i'm not entirely sure how i can use him optimally to benefit me. i did like him though - i thought he was really neat and i liked how he inspired me to make crazy noises that i would not have otherwise invented.

Although *syngvab* might not be useful for Cara, she can see how it might be beneficial for other types of users:

> Q: Now that you've had this creature in your home for a few days, what do you think of the possibility of more of these types of things being in your personal environment?
>
> A: -if by "these types of things," you mean, "robots that respond to my voice," i think that would be really interesting to have more of them in my personal environment....although i'm not sure how it would personally benefit me. however, i can definitely see voice/sound activated creatures

> helping disabled people (blind, physically handicapped, speech impaired, etc) perform tasks around the house, helping them with every day life. i can also envision these creatures helping individuals in speech therapy programs, helping people with emotional/mental trauma connect with their feelings in a different way, and even helping aspiring vocalists learn how to use their voices and the muscles involved (diaphragm, facial muscles, etc).

Thus *syngvab* has potential use outside of a purely research environment, in the space of non-speech interfaces and as a tool for the handicapped.

⁂ ⁂ ⁂

This case study showed how, in one example, *syngvab* was a participant in the formation of a human-object dyad. Cara was provoked to consider *syngvab* as an *existing* creature, more than just an object, but a creature that had it's own "personality traits" and for which she wanted to see how it moved. We were able to see the trace of the interactions with the creature in the resposnes of Cara, and observe how they changed over the course of the interaction.

So now that we have explored in great depth the constituent aspects of this thesis, it is time to take stock of where we have gone and what possibilities might now exist. As we conclude, we will see how two parts of this thesis, creatures and non-speech sounds, can be split apart and used independently on their own terms in projects. As well, we will see what new spaces for work are presented to us as a result of this thesis.

9

Possibilities That Now Exist

Throughout this thesis I have tried to pull from a variety of disparate areas and domains: science studies, critical theories of the voice, the role of objects in our lives, puppetry and performance studies, artificial intelligence and machine learning, and electronics and motor control. Each area has contributed in varying ways to my thinking, ebbing as the time has progressed, being more of an influence here, less there, yet all being part of the process of the design and study of *syngvab*—less a result, an endpoint, a finality, and more of an opening, an illumination of a new space (or an existing one?) of intriguing possibilities. Thus my decision to name this final chapter in the thesis "Possibilities That Now Exist". This is not a conclusion, a stopping point: the momentum is there for pushing forward, for finding out where my bizarre combination of things leads me, for how it might influence me and possibly change the lives of others.

So to begin with the future, and two projects, one at a later stage than the other. These projects pull from the two main threads that can be seen in the design of *syngvab*. On the one hand, I hope to have convinced you of the power of using creatures, and specifically the non-anthropomorphic, non-zoomorphic kind, of using creatures in playful ways open to interpretation, for drawing people in to new experiences. The first project, *policrae and demochi* , comes from this point of view. The second project, currently untitled, involves working with returning veterans of the Iraq war.

Many veterans have much difficulty transitioning to civilian life, especially because of powerful sonic experiences that occurred while on duty in Iraq and are consequently triggered when they return home. This project focuses on how to create a comfortable space for veterans to respond to these experiences and perhaps communicate them in some way, other than words, to their loved ones. Thus the other side of the coin: the extreme potential that seems to be part of expressing oneself using non-speech sounds.

9.1 *POLICRAE AND DEMOCHI*

I am deeply concerned about the state of political discourse in our country. I fear that we find ourselves in an Us versus Them mentality. Blue state versus Red state. "You're either with us or against us." Political discourse in the contemporary United States seems to be expressed through binary oppositions. Gone is the proverbial grey area of the all-too-recent past, replaced instead by the hard edges, sharp corners that cut those who try and sit on them. To be in the middle is verboten (Kornblut 2006).

Or so we believe.

My contention with *policrae and demochi* is that these diametric oppositions are simply a construction. Rather, I see people as much more nuanced in their views, especially when confronted on a face-to-face basis. When you remove someone from the straight-jacket of 24-hour news channels, limited political parties, and other totalizing groups, I believe you will find at least a kernel of reason. Through addressing a person's own beliefs, their person thoughts, rather than those of their supposed affiliation, we grab a thread with which we can knit a conversation. And it is this thread of reason, this ability to see shades of grey, that I am addressing with *policrae and demochi* .

Current political discourse turns the person holding the opposing view into the Other, a stranger. Thus we are now all strangers in our own land, unable to speak to those with whom we disagree. We migrate to safe groups of friends who do not challenge our views, we read, post, and comment on partisan blogs: in sum, we avoid confronting the stranger. Yet we fail to realize that indeed, in the words of Kristeva, the stranger is within us. "The foreigner is within me, hence we are all foreigners." By recognizing the foreigner, the stranger in us, we gain the ability to be in the place of the Other. "It is not simply—humanistically—a matter of our being able to accept the other, but of being in his place, and this means to imagine and make oneself other for oneself" (Kristeva 1994, p. 13). The goal is not to overcome, to defeat the stranger (for that would mean defeating ourselves), but rather to confront it, to recognize the strangeness in ourselves and in each other.

Yet such a process of recognition often needs a catalyst, and it is this catalyst I

Figure 9.1: Images of *policrae and demochi* . *Top*: model. *Bottom*: installation.

tried to design through *policrae and demochi* [1]. Zoomorphic, the creatures present themselves with only the ability to speak, not the ability to hear (Figure 9.1). They sit passive on a bench, repeating incendiary commentary from both sides of the political spectrum, pulled from contemporary blog postings on liberal and conservative blogs (Figure 9.2). As well, the creatures insert their own thoughts about what they "read", highlighting the ways in which each side objectifies the other.

The aim of the creatures is to provoke passers-by into the twin activities of reflection and conversation. They are a transitional object in the words of Winnicott, creating a potential space of experience. Unlike conversations between two humans, the observation of creatures talking creates a Freudian uncanny experience, decontextualizing the words in multiple ways. First, the unavoidable observation of non-human, non-obviously-animal creatures requires special attention and cannot be easily ignored like humans all-too-often are. Second, the speaking of words through mechanical voices, removing them from the comfort of the glowing computer screen, allows us to see the words for what they are: sounds constructing a barrier between ourselves and others. Upon hearing the words, and seeing the responses of those surrounding the creatures, people will be forced into confronting the sounds around them. Through talking to others about the creatures, they will at the same time talk about the content of the creatures' discussion. This indirect discourse will provide a mediating effect, softening otherwise brutal words. Discussion becomes less about the other person, and more about the creatures. People project onto the creatures psychological issues. Consensus recedes as a goal[2]. The evolution of this process encourages personal discourse with people seen as the Other.

I take these creatures to be cyborgs in the terminology of Donna Haraway: "The main trouble with cyborgs, of course, is that they are the illegitimate offspring of militarism and patriarchal capitalism, not to mention state socialism. But illegitimate offspring are often exceedingly unfaithful to their origins" (Haraway 1991, p. 151). The creatures are not placid, subservient to the control of the capitalist system. They provoke the humans around them. Indeed, by creating these creatures, I call upon the playful nature of cyborgs, aligning myself with others who see emancipatory possibilities in technology-art combinations, rather than divestment from practices that are admittedly aligned with military, government, and corporate interests. "From another perspective, a cyborg world might be about lived social and bodily realities in which people are not afraid of their joint kinship with animals and machines, not

[1] "policrae" from *polis*, and demochi, from *demos*.

[2] "To believe that a final resolution of conflicts is eventually possible—even if it is seen as an asymptotic approach to the regulative idea of rational consensus—far from providing the necessary horizon of the democratic project, is something that puts it at risk. Indeed, such an illusion carries implicitly the desire for a reconciled society where pluralism would have been superseded" (Mouffe 2000, p. 32).

Figure 9.2: Screenshots of interactions with *policrae and demochi* installation at the Stata Center at MIT.

afraid of permanently partial identities and contradictory standpoints" (Haraway 1991, p. 154).

As seen in Figure 9.2, I premiered this work in the MIT Stata center in December, 2006. The results were not as I had hoped, for many fairly obvious reasons. The first was timing: the work went up during finals week at MIT, on one of the coldest days of the year, a combination that is simply deadly: the last thing people wanted to do was to stop their busy schedules and talk amongst themselves about things other than problem sets and finals. I had originally planned to present the creatures at the Boston Common in a location bound to have lots of daily foot traffic, especially over the lunch hour. My thought was that such a location would encourage people to stop and try to figure out what was going on, an action that would hopefully get them to start to speak with one another. This leads to the second reason why the project was not as much of a success as I had hoped: because of logistical problems I had to place to creatures on a bench near the wall, preventing people from gathering around. The space was closed, not as inviting as I had desired. With the ability to walk around the creatures, there would be more of a possibility for short-term, spontaneous engagement. The creatures would be seen less as a sculpture, meant to be looked at, and more as a part of the environment, something to be discussed. Finally, the creatures were not "responsive enough" to the environment. They followed a script, much like the course of political discourse. Yet people expected some sort of reaction by the creatures to their (the people's) presence. As well, for the acoustics of the space the creatures' voices were entirely too soft; while this meant that people had to get close to the creatures to hear what they were saying, this meant that many passed-by, thinking *policrae and demochi* were simply another MIT prank.

9.2 NON-SPEECH AND THE VETERAN

What is especially chilling about this war in Iraq is the magnitude of the difficulty faced by soldiers upon their return. There is no training for return to civilian life. While we train the soldier for the challenges of the battlefield, we do not train them for the challenges of the home front—and it is indeed a front, for the switch from a combat (or even support) soldier to civilian is not seamless, is not immediate, and does not, cannot, following the logic of a digital switch.

It has been known for some time that soldiers have incredibly fierce, overpowering responses to stimuli upon their return from combat. Anything that might resemble their experience, especially in the aural realm, can trigger a fight-or-flight response—the dropping of a tray in a restaurant, a car backfiring, fireworks. This is not merely a psychological response, "all in the head" as many might want to do to dismiss it. It is also physiological, a natural reaction of the body to stress (Blanchard, Kolb, Pallmeyer, and Gerardi 1982; Orr, Solomon, Peri, Pitman, and Shalev 1997; Bremner, Staib, Kaloupek, Southwick, Soufer, and Charney 1999; Ferguson, Cassaday, and Bibby 2004). It is a symptom of post-traumatic stress disorder (PTSD), a condition that is all-too-often mis-diagnosed or entirely ignored, either by medical professionals or by the soldier himself or herself. Returning to civilian life is difficult enough and becomes even more challenging coupled with the symptoms of PTSD.

There is an interesting trend, nevertheless, of using technology, specifically virtual reality, to treat PTSD(Rothbaum, Hodges, Alarcon, Ready, Shahar, Graap, Pair, Hebert, Gotz, Wills, and Baltzell 1999; Mrdeza and Pandzic 2003; Rizzo, Pair, McNerney, Eastlund, Manson, and Buckwalter 2005; Rizzo, Pair, Graap, McNerney, Wiederhold, Wiederhold, and Spira 2006; Pair, Allen, Dautricourt, Treskunov, Liewer, Graap, Reger, and Rizzo 2006). The thought in many circles is that gradual, controlled exposure to situations that resemble the trauamtic experience will enable the person to figure out the most appropriate coping mechanism for himself or herself. While the goal is not to *cure* PTSD *per se*, since the original trauma cannot ever be eliminated, the hope is that there can be a time when the trauma can be *managed* and the person can live her life without constant reference to a past event.

Emboldened by this work, and through a class with in the Visual Arts Program at MIT taught by Krzysztof Wodiczko entitled the "War Veteran Vehicle", I began to talk with Iraq war veterans about their aural experiences upon their return. I was especially interested in how they personally deal with these overwhelming experiences. One veteran I spoke to talked of his experiences in the mall: how the loud crowds would cause him to feel as if he were back in Iraq.

With my collaborator on this project, Monica Haller, we wanted to create a device that would record the person's physiological response around the time of the triggering event: the elevated heartbeat, the heavy breathing, the increased amount of sweat. At the same time, or perhaps at a later point, the person could then respond to this event in some way, perhaps by expressing the strength of the even through a captured scream, or other type of non-speech sound. The person could then share this event and his response to it with his family, friends, or the public at large. Without speech, using only the physiological signals and the person's response to them, people would have to respond in a visceral way, bypassing the cognitive and going straight to the emotional, to the realm of feelings and internal life.

While I write somewhat glibly of the benefits of the project, there are also potentially serious psychological risks. We have been working with therapists in the area to discuss these problems and frame the project in such a way to minimize them. In many ways the interactions with the people in this project are similar to those with *syngvab*: personal, immediate, intense. In this project we are focusing strongly on the power of non-speech, suggesting that its expression might not only be important and interesting, but potentially therapeutic as well.

9.3 WHAT WE HAVE LEARNED

In this thesis I have attempted to draw from a disparate set of influences to create the mongrel object called *syngvab*. Along the way we have seen how the voice is not simply a signifying tool, but can reflect deeply felt, potentially hidden and unexpressed aspects of one's psychological life. The voice can be encouraged to break free of the confines of language, liberating the unspeakable for one's benefit. We've seen how objects do not merely remain passive, but are actively part of our lives, influencing our actions and activities. These objects play a strong role as relationship partners, especially in the case of relational artifacts, with the concern of dependence, but also concomitant potential for novel experiences. But if the objects we are creating are really new, then perhaps they should be studied in an innovative way as well. We've gone through the arguments for a new conception of technology and its study that does not merely reflect existing assumptions that are incongruent with present-day technical realities and experiences. And we've gone through the use of puppets as a means for embodying these somewhat radical ideas.

We saw that it was possible, through a number of design iterations, to create a non-anthropomorphic, non-zoomorphic creature and to provide an engaging, inviting home (stage) for it. I have shown that it is possible to discriminate between human speech and non-speech with a small number of features and with high accuracy. This enabled the development of a software agent that would respond in real-time to the non-speech sounds of a person. We saw through a case study that this interaction with the creature left a strong impression on the participant, suggesting the strength of not only the expression of non-speech sounds, but also the possible benefits of interacting with a non-representational robotic creature.

9.4 WHAT IS LEFT TO DO

The sagacious reader of this thesis will already have an idea of things upon which I could improve, but for my benefit I want to go through a number of them here.

- *Better discrimination between speech and non-speech.* The SVM classifier I developed works remarkably well at correctly choosing whether a frame is speech or non-speech. Taking only four relatively generic features over a small subset of subjects and recording conditions enabled great separation. However, I could do better. Over a set of twenty, thirty features or more I could take a statistical distribution of measurements for speech and non-speech. From here I could select the subset of features that have the least variance within-class, while at the same time having the largest variance between-class. This should give me a set that would be most likely to be useful in discriminating between the two classes. Additional samples of speech and non-speech, from more speakers, would be helpful as well.

- *Better calculation of limits for scaling.* At the moment I have a subjectively calculated set of limits for my scaling parameters in the agent. A better option would be an adaptive system, perhaps using self-organizing maps as in Downie (2005, pp. 145–148).

- *Mechanical improvements.* The current marionette system works quite well in many cases, especially scripted movements. However, there are definite problems when we let the agent move the marionette at will. This is due to the fact that there is no encoded knowledge of the relationship between the strings of the creature. This causes strings to get tangled on the pulleys, wrapping around the spoke of the motor instead of the pulley. This could be remedied in one of two ways: either encode the relative positions of the strings, preventing any movement that would cause tangling; or design a better string guidance system that would prevent physical tangling.

- *Consider swing dynamics.* At the moment there is no consideration of the swing dynamics of *syngvab*, which are great due to the fact that the object is not rooted to the ground. There is some work by Filipic, Urbancic, and Krizman (1999) that uses a genetic algorithm approach to creating a control system to decrease the undesirable swinging motions. This could be coupled with the next point to allow as much dynamic swinging motion as desired.

- *Larger number of participants to 'guide' behavior.* For this thesis I was the only person to rate and give fitness values during the interactive evolutionary algorithm. This aspect of the motor control would undoubtedly be improved with additional participants to help choose more interesting movement possibilities.

- *More participants in the in-situ studies.* Because of time constraints and the requirement for semi-long-term interaction, I was only able to put *syngvab* in the home of a few users. I would like to greatly increase this number to see the variety of interactions. Is there anything about age, gender, background that seems to correlate with responses? How might the interactions differ on a one-week, one-month, even one-year timescale?

- *Effect of my position on the response to syngvab.* One thing not described to this point is the role I played in the introduction of *syngvab* to participants. In what ways does my position cause someone to immediately ascribe certain personality traits to the creature? A certain gender? How does my position as someone from MIT affect the way people respond to the creature. I would like to study this by hiring an actor to introduce *syngvab* to participants. This would be novel and should help us to better understand the parts we play as members of the technocracy in our interactions with the public.

- *Self-reflection on certain qualities of syngvab.* I spent a large amount of time in Chapter 4 detailing how technological systems reflect certain aspects of society, internalizing assumptions that may not be appropriate. While I have tried to be aware of this during the development of the thesis, I am certain there are choices I made that contradict this desire. With some distance from the thesis and the work I should be able to better understand my role as an attempted outsider (but still member of) a particular technological system.

9.5 WHAT WE HAVE LEFT TO EXPLORE

The most serendipitous discovery during the thesis was the powerful nature of making and hearing non-speech sounds. More than the classifier, more than the simulation, more than the creation of a robotic marionette system—non-speech was a potent means of expression. It is especially interesting to consider how non-speech lives outside of codified technological norms that increasingly try to delimit what is acceptable or not.

Yet there is much left to explore in this area, now that we have identified the intense nature of this unspoken part of our lives. As suggested by not only Cara in Chapter 8, but also others with whom I spoke during the thesis, non-speech sounds are absolutely important for whole groups of people often ignored by technological development, especially the handicapped. Instead of focusing on the instrumental *wants* of a distant designer, we can focus on the actual *needs* of the person (Papanek 1985). To take one example, for many low-functioning autistic children speech is difficult if not impossible. They make sounds that are maybe close to speech but do not pass as phonemes; they remain outside of language. An object like *syngvab* might enable the child or adult to learn different ways of expressing themselves; or *syngvab* might be an object that they could use to communicate to others, even if for some users, like Cara , non-speech is not often used for communication. Additionally, in discussions with drama therapists I learned of the use of non-speech by actors in the process of learning lines for plays. Instead of endlessly repeating the words of the playwright, perhaps stripping all meaning away from them, the actor makes sounds that reflect the *prosodic* content of the text. Thus in this case, non-speech brings the actor closer to the text and its encoded emotional content by abstracting away the codified norms of language.

Additionally, there is much to explore with non-representational robotics. We have seen how the creation of the *syngvab* creature encouraged an active meaning-making process on the part of the participant. Since *syngvab* did not immediately resemble any existing creature, the participant had to involve herself in the creation of her own story for the creature, her own understanding of its personality. I believe that this, coupled with the power of puppetry described in Chapter 5, is sorely underutilized in many areas of HCI. We would do well as a field to draw from this tradition as a means of breaking down the barrier between object and human through the sometimes humorous, but still serious, play of the puppet and creature.

⁂ ⁂ ⁂

This thesis has been a journey of my own to traverse a host of intellectual boundaries: critical theory, sociology, HCI, electronics, artificial intelligence. The progress made along the way resulted not in a synthesis, but a mongrel, a goo shall we say, strands of thought intersecting in unexpected ways. My experience suggests that this is the future: as we better understand our folly of needlessly splitting disciplines apart and forcing them to exist in separate fiefdoms, we discover how everything was interrelated in the first place. It is not our goal to separate and categorize, to blindly assume that nature should live in tidy boxes. Rather, our desire can be to see where the connections between disparate thoughts are, how we can use those links to think new things. We can look at the negative spaces and observe what is not said, what is not known, what is not thought. In the end there is no end, only an opening of possibility, an expansion of the space of existence.

A

Motor Control Boards

This appendix includes the schematics and board layout for the motor control system, with both the master control and stepper motor control boards:

1. Figure A.1: Schematric of the Master Control circuit
2. Figure A.2: Board outline of the Master Control circuit
3. Figure A.3: Schematric of the Stepper Motor Control circuit
4. Figure A.4: Board outline of the Stepper Motor Control circuit

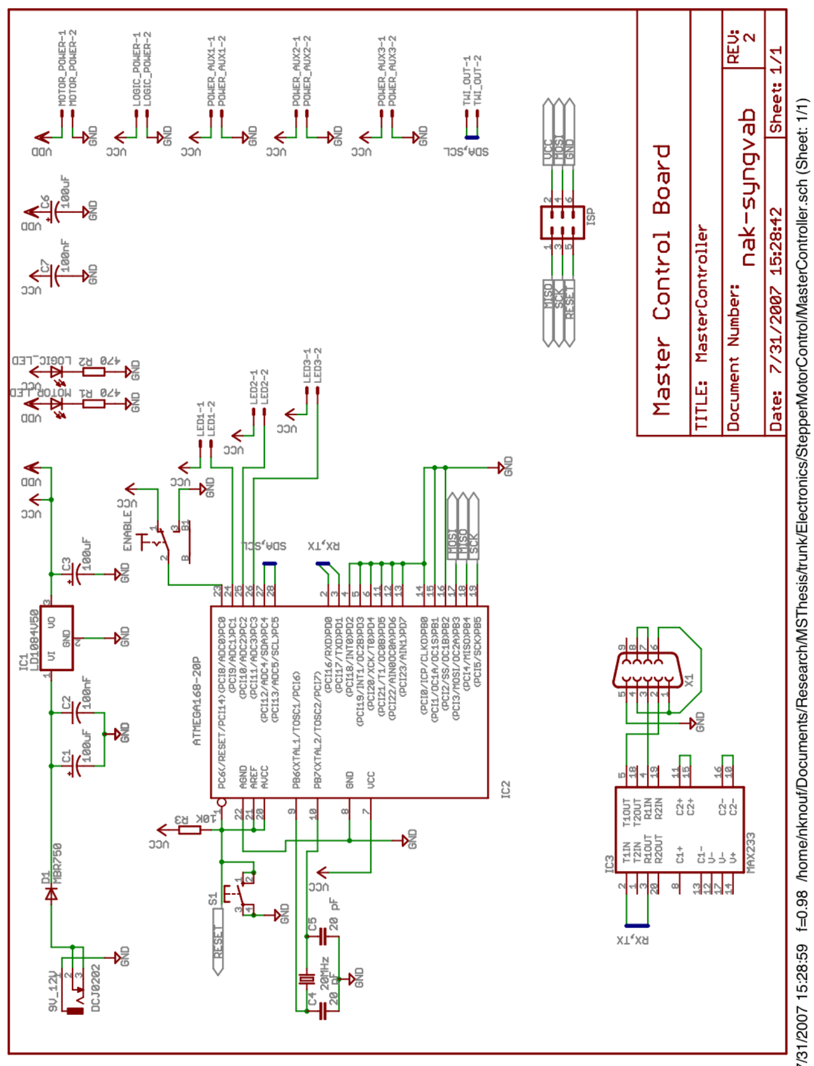

Figure A.1: Schematic of Master Control Circuit

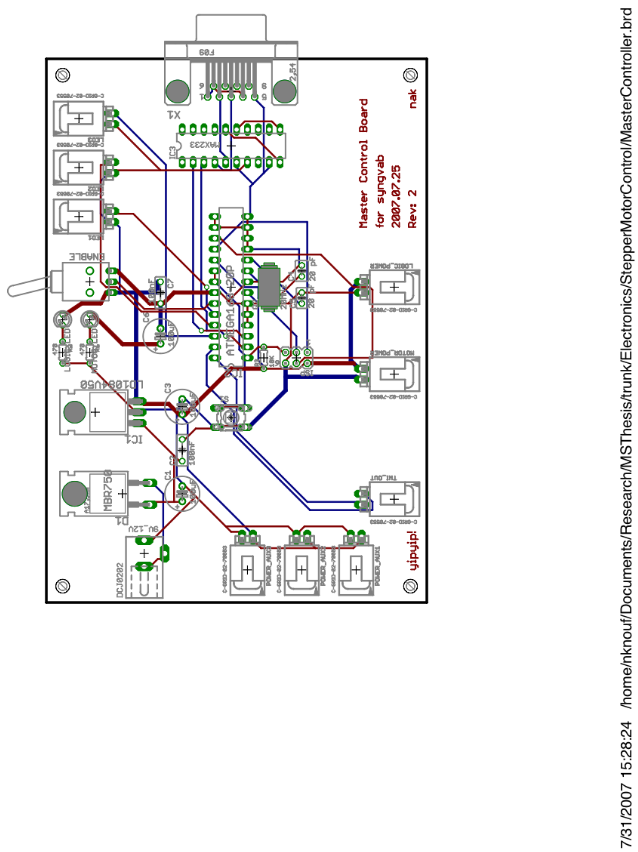

Figure A.2: Board Outline of Master Control Circuit

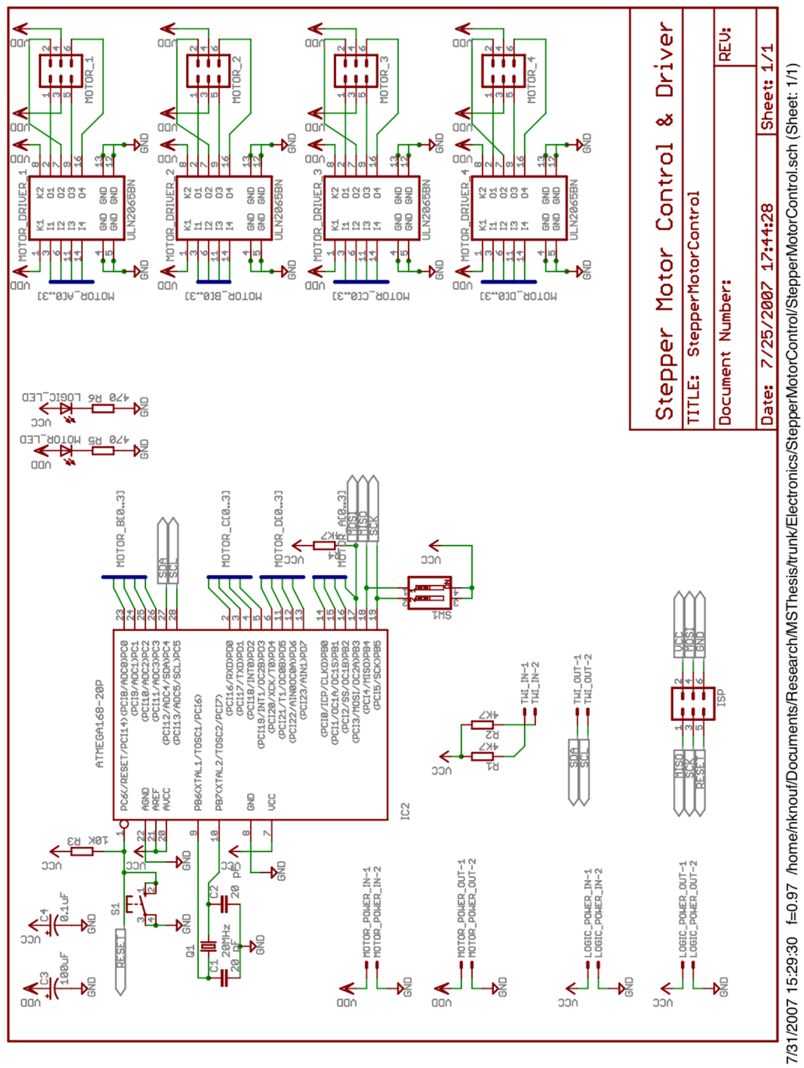

Figure A.3: Schematic of Stepper Motor Control Circuit

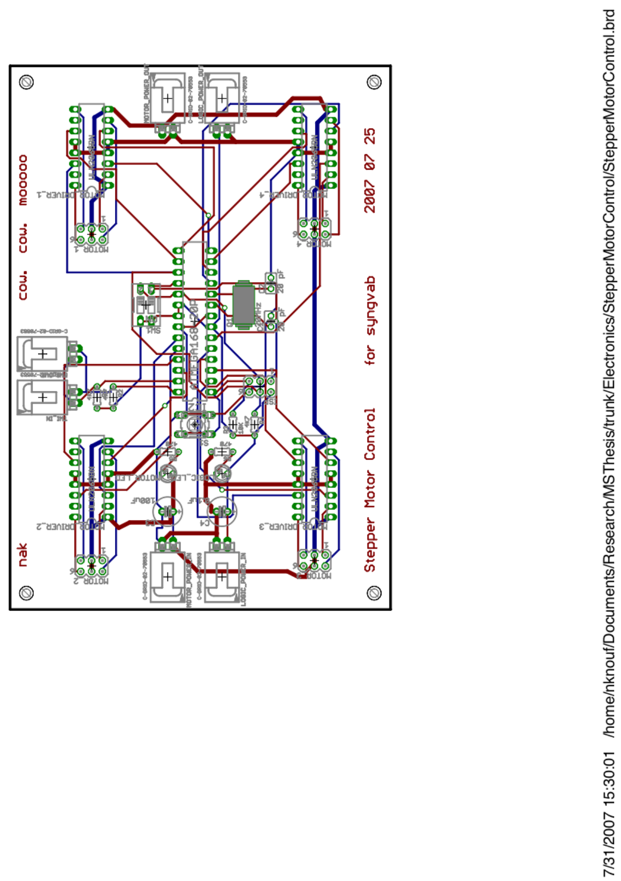

Figure A.4: Board Outline of Stepper Motor Control Circuit

B

Stage Design

This appendix includes diagrams for important parts of the stage. All dimensions are in inches.

Briefly, there are two stage platforms, one lower and one upper (Figure B.1). There are two stage side supports, one for the left and one for the right. There are five one-inch square solid aluminum bars that form the vertical supports. These attach to the stage platform through L-brackets; you will have to figure out the correct placement of drill holes for your given L-bracket. The stage left and right side supports attach to the outside of the aluminum bars on the left and right sides.

Attached to the top of the top stage platform are the mounting brackets for the motors (Figure B.3). Drill holes for attachment to the top stage platform must be decided depending on your desired location for the motors. Of course, the bracket design will have to be modified depending on your choice of motor.

B. STAGE DESIGN

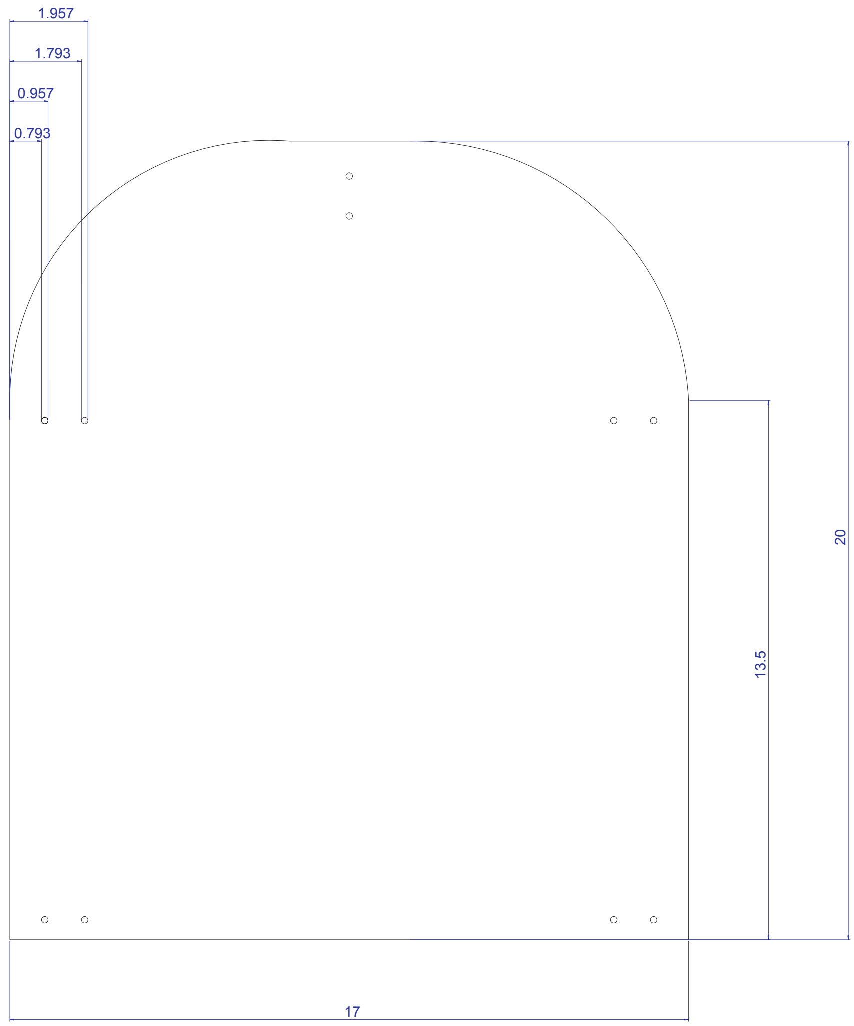

Figure B.1: Stage Platform Diagram

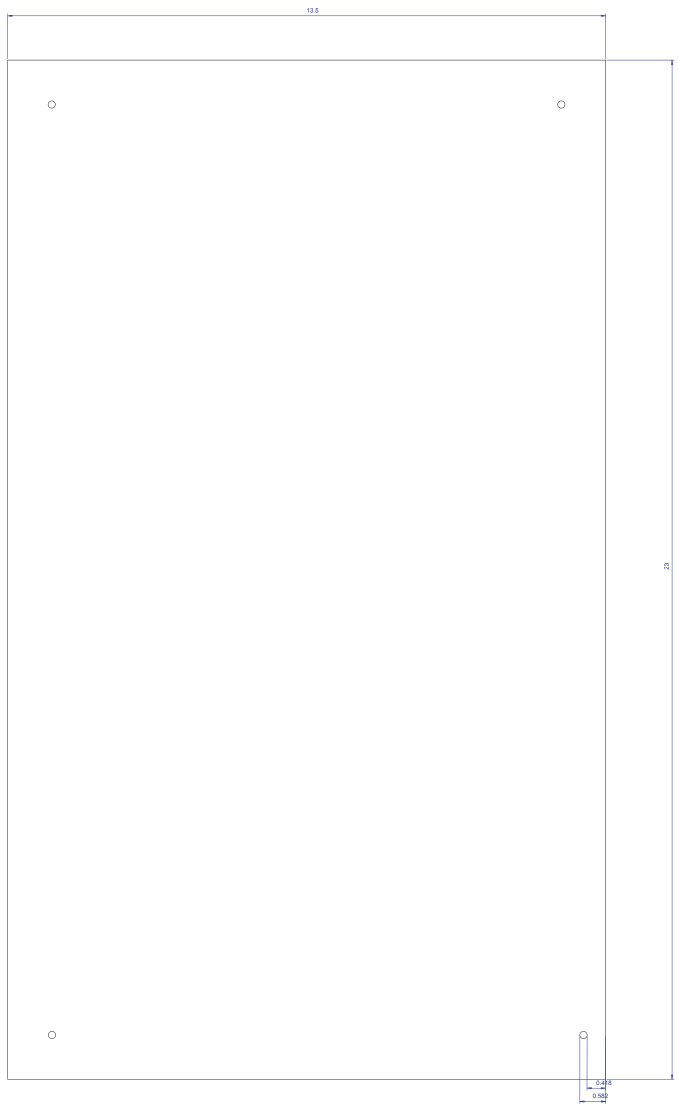

Figure B.2: Stage Left and Right Side Supports

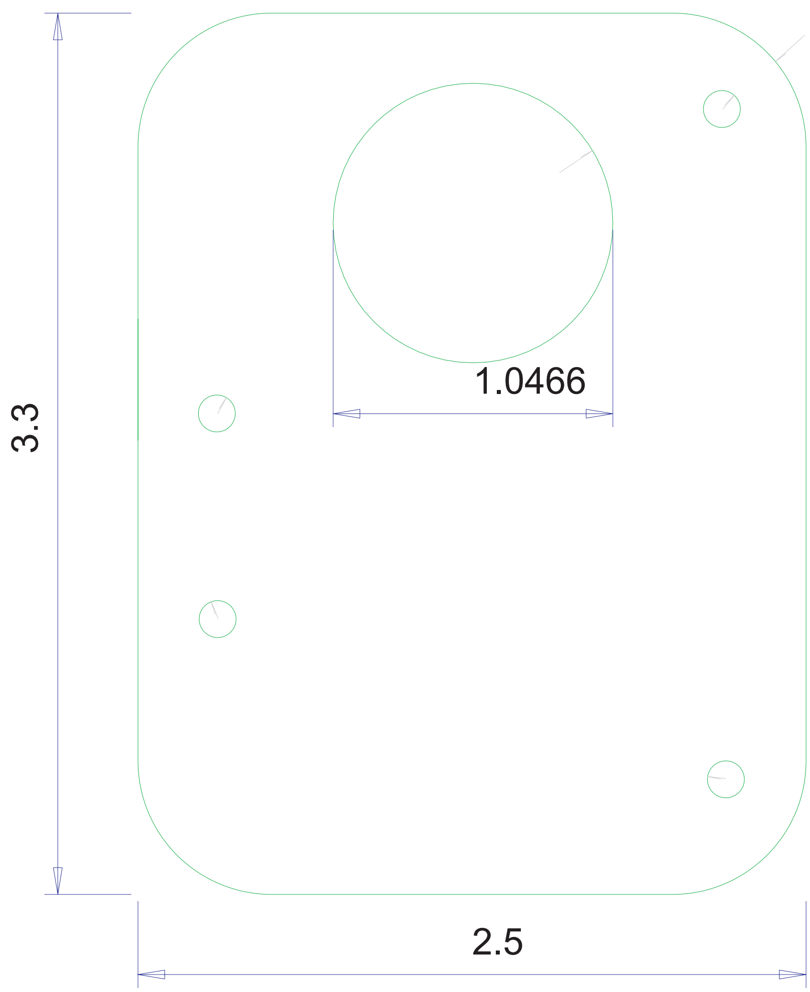

Figure B.3: Motor Mounting Brackets

C

Consent Forms

All *in-situ* research was approved by the Committee on the Use of Humans as Experimental Subjects (COUHES) at MIT. Included here are the consent forms used for this research. Note that I did not perform the study described in the first form, the evaluation of movements in the simulated creature.

C.1 BEHAVIORAL EXPERIMENT WITH A ROBOTIC CREATURE SIMULATION

CONSENT TO PARTICIPATE IN NON-BIOMEDICAL RESEARCH

Human Interactions with Robotic Creatures

Experiment 1: Behavioral Experiment with a Robotic Creature Simulation

You are asked to participate in a research study conducted by Nicholas Knouf, B.S. and Tod Machover, B.M., M.M., from the MIT Media Lab at the Massachusetts Institute of Technology (M.I.T.). Results from this study will be included in Nicholas KnoufÃ•s MasterÃ•s thesis. You were selected as a possible participant in this study because you volunteered. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

PURPOSE OF THE STUDY

The purpose of this study is to investigate human perception of motions of a simulated robotic creature.

PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things:

First, we will explain the purpose of the study and give a brief tutorial on how to respond using our computer program.

Then we will show you a series of motions of a simulated robotic creature. Following each motion we will ask you a few questions about your perception of what you have just seen. Throughout the study, you are encouraged to take as much time as you need, and ask for clarification along the way.

You are encouraged to ask questions at any time during the study if you are unsure of the meaning of particular questions or if you are having difficulty with the computer program.

The total time for this experiment is one half-hour and will take place entirely in the laboratory.

POTENTIAL RISKS AND DISCOMFORTS

We do not anticipate any physical or psychological risks as a result of your participation in this study.

POTENTIAL BENEFITS

We do not anticipate that you will benefit in any significant way as a result of your participation in this study, beyond the remuneration described below.

The benefits for society are a better understanding of effective and engaging robotic creature movement.

PAYMENT FOR PARTICIPATION

You will receive $5 for participation in this experiment, expected to last one half-hour. If for any reason you withdraw from the study, you will be paid a prorated amount.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law.

Information associated with your participation in the study will be identified by an ID that has no connection with your personal information. As well, this information will be stored on a secure computer to which only the experimenters have access.

IDENTIFICATION OF INVESTIGATORS

If you have any questions or concerns about the research, please feel free to contact either of the investigators: Nicholas Knouf (nknouf@media.mit.edu, 617-452-5633) or Tod Machover (tod@media.mit.edu, 617-253-1613).

EMERGENCY CARE AND COMPENSATION FOR INJURY

"In the unlikely event of physical injury resulting from participation in this research you may receive medical treatment from the M.I.T. Medical Department, including emergency treatment and follow-up care as needed. Your insurance carrier may be billed for the cost of such treatment. M.I.T. does not provide any other form of compensation for injury. Moreover, in either providing or making such medical care available it does not imply the injury is the fault of the investigator. Further information may be obtained by calling the MIT Insurance and Legal Affairs Office at 1-617-253 2822."

RIGHTS OF RESEARCH SUBJECTS

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.

C.2 IN-SITU EXPERIMENT WITH A ROBOTIC CREATURE

CONSENT TO PARTICIPATE IN NON-BIOMEDICAL RESEARCH

Human Interactions with Robotic Creatures

Experiment 2: In-situ Experiment with a Robotic Creature

You are asked to participate in a research study conducted by Nicholas Knouf, B.S. and Tod Machover, B.M., M.M., from the MIT Media Lab at the Massachusetts Institute of Technology (M.I.T.). Results from this study will be included in Nicholas Knouf's Master's thesis. You were selected as a possible participant in this study because you volunteered. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate.

PARTICIPATION AND WITHDRAWAL

Your participation in this study is completely voluntary and you are free to choose whether to be in it or not. If you choose to be in this study, you may subsequently withdraw from it at any time without penalty or consequences of any kind. The investigator may withdraw you from this research if circumstances arise which warrant doing so.

PURPOSE OF THE STUDY

The purpose of this study is to investigate human experiences with a robotic creature in a home environment as well as means of encouraging new types of personal expression.

PROCEDURES

If you volunteer to participate in this study, we would ask you to do the following things:

First, we will explain the purpose of the study and conduct a short interview as to your experiences and thoughts of robotic creatures.

Next, we will schedule a time to visit your home, bringing the robotic creature with us. In your home we will, in consultation with you, determine the most appropriate location for the creature and its equipment. We will explain basic interaction with the creature and use of the equipment as well as minimal troubleshooting. (We will also give you detailed contact information in case of unexpected problems.) Following this tutorial we will explain the use of an on-line diary system.

For the next week we will ask you to interact with the robotic creature for a total of at least one half-hour per day. We will also ask you to contribute to the on-line diary at least once per day.

At the end of the week we will return to your home at a mutually agreeable time. We will remove the robotic creature and its associated equipment. We will then conduct a short interview as to your experiences. At this time we will schedule a follow-up interview to be conducted two weeks later in person or over the phone.

You are encouraged to ask questions at any time during the study.

POTENTIAL RISKS AND DISCOMFORTS

We do not anticipate any physical or psychological risks as a result of your participation in this study. You may experience mild embarrassment while exploring expressive possibilities with the creature.

POTENTIAL BENEFITS

We do not anticipate that you will benefit in any significant way as a result of your participation in this study. However, your participation may enable you to explore new means of personal expression.

The benefits for society are a better understanding of how people interact with robotic creatures, as well as ways to encourage personal expression.

PAYMENT FOR PARTICIPATION

You will receive no payment for participation in this study.

CONFIDENTIALITY

Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law.

Information associated with your participation in the study will be identified by an ID that has no connection with your personal information. Your online diary entries will be encrypted in transit using industry-standard techniques. As well, this information will be stored on a secure computer to which only the experimenters have access.

We would like to audiotape your three interviews: prior to bringing the creature to your home, immediately after the one-week period of having the creature in your home, and during the follow-up interview two weeks later.

Unless you give us permission to use your name, title, and / or quote you in any publications that may result from this research, the information you tell us will be confidential.

We would like to record these interviews on audio cassette so that we can use it for reference while proceeding with this study. We will not record these interviews without your permission. If you do grant permission for this conversation to be recorded on cassette, you have the right to revoke recording permission and/or end the interview at any time.

This project will be completed by February 28th, 2008. All interview recordings will be stored in a secure workspace until one (1) year after that date. The tapes will then be destroyed.

(Please check all that apply)

[ ] I give permission for this interview to be recorded on audio cassette.

[ ] I give permission for the following information to be included in publications resulting from this study:

[ ] my name

[ ] my title

[ ] direct quotes from this interview

IDENTIFICATION OF INVESTIGATORS

If you have any questions or concerns about the research, please feel free to contact either of the investigators: Nicholas Knouf (nknouf@media.mit.edu, 617-452-5633 (daytime), 617-388-0567 (evening)) or Tod Machover (tod@media.mit.edu, 617-253-1613).

EMERGENCY CARE AND COMPENSATION FOR INJURY

"In the unlikely event of physical injury resulting from participation in this research you may receive medical treatment from the M.I.T. Medical Department, including emergency treatment and follow-up care as needed. Your insurance carrier may be billed for the cost of such treatment. M.I.T. does not provide any other form of compensation for injury. Moreover, in either providing or making such medical care available it does not imply the injury is the fault of the investigator. Further information may be obtained by calling the MIT Insurance and Legal Affairs Office at 1-617-253 2822."

RIGHTS OF RESEARCH SUBJECTS

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.

Glossary

actor-network theory (ANT) A methodology for the study of technical and scientific systems that works to consider the role of actors (humans and nonhumans) in the *social* construction of knowledge., 56, 57

degrees of freedom (DoF) Number of possible parts of a system that can be moved, either by an external force or through the actions of gravity. A physical system that has a number of degrees of freedom of control less than the total number of degrees of freedom is said to be *underarticulated*. In most cases, marionettes are underarticulated systems, with the consequent result that the rest of the bodies (joints and parts of limbs) are under the influence of gravity., 65, 69

echo-state network (ESN) Type of recurrent network that allows for easy training by only needing to train the output weights of the network., 112, 114, 115, 117

Feed-forward Neural Network (FFNN) Type of neural network where links between nodes are limited to only the forward direction. Information "flows" from the input units, through the units in the hidden layer, and finally through the output units. These networks are the most common type of neural net and have found wide application in classification, generalization, and prediction tasks., 111

Human-Computer Interaction (HCI) Discipline that studies the ways in which people interact with computers and technological objects. HCI is decidedly cross-disciplinary, drawing from computer science, cognitive science, anthropology, and the arts., 121, 138, 139

Mel-frequency cepstral coefficients (MFCC) Set of critical band filters that approximate the response of the human inner ear., 106

Open Dynamics Engine (ODE) Open-source rigid body physics simulator, 94, 96

Open Sound Control (OSC) A network-based (UDP) protocol for communicating amongst computers, synthesizers, and other multimedia devices. The protocol is general enough to be used as a easy means of sending data between different programs and processes, as there are many open-source libraries for common programming languages and operating systems., 97, 118

pure data (PD) An open-source, real-time audio processing environment, similar in design to Max/MSP., 75

Proportional-Integral-Derivative controller (PID) Method for calculating the error in a measurement by considering the difference, integral, and derivative of the measured value as compared to some setpoint. Often used in control applications., 96

post-traumatic stress disorder (PTSD) A condition that involves intense physiological and psychological responses to a prior trauma. The responses can be triggered by stimuli that resemble aspects of the trauma in some way. The condition can last for years or decades after the trauma has occurred. PTSD is especially prevalent in returning war veterans and is woefully underdiagnosed and untreated., 134

persistent uniform resource locator (PURL) A means of creating persistent links to networked resources through the use of an intermediate resolution service. The PURLs used in this thesis are provided by the Online Computer Library Center (`http://purl.org`)., 18

Pulse-Width Modulation (PWM) Controlling the effective amplitude (or, in electronics, voltage) of a signal by changing the amount of time (*duty cycle*) the signal is high. Commonly used to control the speed of DC motors., 88

radial-basis function (RBF) Kernel function commonly used with support-vector machine classifiers., 107

Recurrent Neural Network (RNN) Type of neural network where links between nodes are not limited to only the forward direction; links can be from higher layers to lower layers, or from one node to itself. This enables the representation of temporal information, with the consequent difficulty of training and network design., 110–112

RObotic Marionette System (ROMS) Robotic marionette project from the work of I-Ming Chen and colleagues at the Nanyang Technological University in Singapore. ROMS is a full 16-DoF system enabling both smooth and gross movement of a human character. The direction of the project is influenced by traditional Chinese Gou Pai marionette theater., 68, 69

Inter-Integrated Circuit (I2C) Chip-to-chip communication protocol over two wires; the name is trademarked by Phillips, hence the similar protocol named Two-Wire Interface (TWI)., 89

short-time Fourier transform (STFT) A form of the Fourier transform used to determine short-term, transient changes in frequency content of a discrete signal. In the STFT there is a trade-off between frequency and temporal resolution., 105, 106

Support Vector Machines (SVM) Type of machine learning technique commonly used for classification. Extremely robost and tolerant of non-separable data., 107, 136

Two-Wire Interface (TWI) Chip-to-chip communication protocol requiring, as its name suggests, only two lines. See also I2C., 89, 90

Bibliography

Ackermann, Edith K. (1991). "The Agency Model of Transactions: Toward an Understanding of Children's Theory of Control". In: *Constructionism Part 4: Cybernetics and Constructionism*. Ed. by I. Harel and S. Papert. Norwood, New Jersey, USA: Ablex Publishing Corporation. 367–379.

Agre, Philip E. (1997). "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI". In: *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*. Ed. by Geoffrey C. Bowker, Susan Leigh Star, William Turner, and Les Gasser. Lawrence Erlbaum Associates, Inc. 131–157.

Angeline, P. J., G. M. Saunders, and J. B. Pollack (1994). "An evolutionary algorithm that constructs recurrent neural networks". In: *Neural Networks, IEEE Transactions on* 5.1. 54–65.

Artaud, Antonin (1958). *The Theater and its Double*. New York, NY, USA: Grove Press.

Austin, David (2004). *Generate stepper-motor speed profiles in real time*. 2004. URL: http://www.embedded.com/showArticle.jhtml?articleID=56800129 (visited on 12/30/2004).

Baldassarre, Gianluca, Stefano Nolfi, and Domenico Parisi (2003). "Evolving Mobile Robots Able to Display Collective Behaviors". In: *Artificial Life* 9.3. 255–267. DOI: 10.1162/106454603322392460. URL: http://www.mitpressjournals.org/doi/abs/10.1162/106454603322392460.

Bell, John (2001*a*). "Puppets, Masks, and Performing Objects at the End of the Century". In: *Puppets, Masks, and Performing Objects*. Ed. by John Bell. TDR Books. MIT Press.

— (2001*b*). "The End of *Our Domestic Resurrection Circus*: Bread and Puppet Theater and Counterculture Performance in the 1990s". In: *Puppets, Masks, and Performing Objects*. Ed. by John Bell. TDR Books. MIT Press.

Berenzweig, Adam, Beth Logan, Daniel P.W. Ellis, and Brian Whitman (2004). "A Large-Scale Evaluation of Acoustic and Subjective Music-Similarity Measures". In: *Computer Music Journal* 28.2. 63–76.

Bertenthal, Bennett I (1996). "Origins and early development of perception, action, and representation". In: *Annual Review of Psychology* 47. URL: http://proquest.umi.com/pqdlink?did=9242563&Fmt=7&clientId=5482&RQT=309&VName=PQD.

Biles, J. A. (1999). "Life with GenJam: interacting with a musical IGA". In: *1999 IEEE International Conference on Systems, Man, and Cybernetics*. Vol. 3. 652–656 vol.3.

Blanchard, Edward B., Lawrence C. Kolb, Thomas P. Pallmeyer, and Robert J. Gerardi (1982). "A psychophysiological study of Post Traumatic Stress Disorder in Vietnam veterans". In: *Psychiatric Quarterly* V54.4. 220–229. URL: http://dx.doi.org/10.1007/BF01064817.

Blumberg, Bruce, Marc Downie, Yuri Ivanov, Matt Berlin, Michael Patrick Johnson, and Bill Tomlinson (2002). "Integrated learning for interactive synthetic characters". In: *ACM Trans. Graph.* 21.3. 417–426. ISSN 0730-0301. DOI: http://doi.acm.org/10.1145/566654.566597.

Boehner, Kirsten, Shay David, Joseph 'Jofish' Kaye, and Phoebe Sengers (2005). "Critical Technical Practice as a Methodology for Values in Design". In: *Proceedings of the 2005 CHI Workshop on 'quality value choice'*.

Bogatyrev, Pyotr (2001 [1923]). "Czech Puppet Theatre and Russian Folk Theatre". In: *Puppets, Masks, and Performing Objects*. Ed. by John Bell. TDR Books. MIT Press.

Böhlen, Marc (2006). "When a Machine Picks a Fight: notes on machinic male-dicta and synthetic hissy fits". In: *Proceedings of the 2006 CHI Workshop on Misuse and Abuse of Interactive Technologies*.

Böhlen, Marc, and M. Mateas (2002). "Machines with a different calling [robots for social environments]". In: *Intelligent Robots and System, 2002. IEEE/RSJ International Conference on* 2. 1172–1177 vol.2.

Böhlen, Marc, and JT Rinker (2005). "Experiments with Whistling Machines". In: *Leonardo Music Journal* 15. 45–52.

Bongard, Josh, Victor Zykov, and Hod Lipson (2006). "Resilient Machines Through Continuous Self-Modeling". In: *Science* 314.5802. 1118–1121. DOI: 10.1126/science.1133687. URL: http://www.sciencemag.org/cgi/content/abstract/314/5802/1118.

Borden G. R., IV (2002). "An aural user interface for ubiquitous computing". In: ISBN 1530-0811. 143–144.

Breazeal, Cynthia, Daphna Buchsbaum, Jesse Gray, David Gatenby, and Bruce Blumberg (2005). "Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots". In: *Artificial Life* 11.1-2. 31–62. DOI: 10.1162/1064546053278955. URL: http://www.mitpressjournals.org/doi/abs/10.1162/1064546053278955.

Bremner, J D, L H Staib, D Kaloupek, S M Southwick, R Soufer, and D S Charney (1999). "Neural correlates of exposure to traumatic pictures and sound in Vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study". In: *Biol Psychiatry* 45.7. 806–816. URL: http://www.hubmed.org/display.cgi?uids=10202567.

Cetina, Karin Knorr (1997). "Sociality with Objects: Social Relations in Postsocial Knowledge Societies". In: *Theory, Culture & Society* 14.4. 1–30. URL: http://tcs.sagepub.com.

Chang, Chih-Chung, and Chih-Jen Lin (2007). *LIBSVM: a library for support vector machines*. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chen, C. C. J., and R. Miikkulainen (2001). "Creating melodies with evolving recurrent neural networks". In: *Proceedings of 2001 International Joint Conference on Neural Networks*. Vol. 3. 2241–2246 vol.3.

Chen, I. M., R. Tay, S. Xing, and S. H. Yeo (2004). "Marionette: From traditional manipulation to robotic manipulation". In: *International Symposium on History of Machines and Mechanisms*.

Chen, I-Ming, Shusong Xing, R. Tay, and Song Huat Yeo (2005). "Many strings attached: from conventional to robotic marionette manipulation". In: *Robotics & Automation Magazine, IEEE* 12.1. 59–74. ISSN 1070-9932.

Corebounce (2006). *Corebounce Projects: The Digital Marionette*. 2006. URL: http://www.corebounce.org/wiki/Projects/Marionette (visited on 12/19/2006).

Corporation, Atmel (2006). *AVR446: Linear speed control of stepper motor*. Atmel Corporation.

Dahlstedt, Palle (2001). "Creating and Exploring Huge Parameter Spaces: Interactive Evolution as a Tool for Sound Generation". In: *Proceedings of International Computer Music Conference*. URL: citeseer.ist.psu.edu/699688.html.

Deleuze, Gilles (2006). "Four Propositions on Psychoanalysis". In: *Two Regimes of Madness*. New York, NY, USA: Semiotext(e).

Deleuze, Gilles, and Félix Guattari (1983 [1972]). *Anti-Oedipus: Capitalism and Schizophrenia*. University of Minnesota Press.

DiSalvo, Carl (2004). "The Uncanny and The Everyday in the Design of Robots". In: *Proceedings of the Reflective HCI: Towards a Critical Technical Practice Workshop at CHI 2004*.

Dittrich, Peter, Andreas Burgel, and Wolfgang Banzhaf (1998). "Learning to Move a Robot with Random Morphology". In: *Proceedings of the First European Workshop on Evolutionary Robotics*. Vol. 1468. Paris: Springer-Verlag. ISBN 3-540-64957-3. 165–178. URL: citeseer.ist.psu.edu/dittrich98learning.html.

Dobson, Kelly (2004). "Blendie". In: *DIS '04: Proceedings of the 2004 conference on Designing interactive systems*.

Dobson, Kelly (2000). "BOOM". Master's thesis. Massachusetts Institute of Technology.

— (2005). "Wearable Body Organs: Critical Cognition Becomes (Again) Somatic". In: *Creativity & Cognition 2005*. 259–262.

Dobson, Kelly, Brian Whitman, and Daniel P.W. Ellis (2005). "Learning Auditory Models of Machine Voices". In: *Proceedings of the 2005 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE.

Dolar, Mladen (2006). *A Voice and Nothing More*. Cambridge, MA, USA: MIT Press.

Downie, Marc (2000). "behavior, animation, music: the music and movement of synthetic characters". Master's thesis. Massachusetts Institute of Technology.

— (2005). "Choreographing the Extended Agent: performance graphics for dance theater". PhD thesis. Massachusetts Institute of Technology.

Draves, Scott (2005). "The Electric Sheep Screen-Saver: A Case Study in Aesthetic Evolution". In: *EvoWorkshops 2005*. Springer. 458–467. URL: http://www.springerlink.com/content/ldw3djcxw95lyljh.

Dunne, Anthony, and Fiona Raby (2002). *Design Noir: The Secret Life of Electronic Objects*. Basel, Switzerland: Birkhäuser.

Ellis, Daniel, Brian Whitman, Adam Berenzweig, and Steve Lawrence (2002). "The Quest For Ground Truth in Musical Artist Similarity". In: *Proceedings of the 3rd International Conference on Music Information Retrieval*. Paris, France.

Feenberg, Andrew (1999). *Questioning Technology*. London, UK: Routledge.

Fell, Harriet, Cynthia Cress, Joel MacAuslan, and Linda Ferrier (2004). "visiBabble for reinforcement of early vocalization". In: *Assets '04: Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility*. Atlanta, GA, USA: ACM Press. ISBN 1-58113-911-X. DOI: http://doi.acm.org/10.1145/1028630.1028659. 161–168.

Ferguson, E, H J Cassaday, and P A Bibby (2004). "Odors and sounds as triggers for medically unexplained symptoms: a fixed-occasion diary study of Gulf War veterans". In: *Ann Behav Med* 27.3. 205–214. DOI: 10.1207/s15324796abm2703_8. URL: http://www.hubmed.org/display.cgi?uids=15184096.

Filipic, Bogdan, Tanja Urbancic, and Viljem Krizman (1999). "A combined machine learning and genetic algorithm approach to controller design". In: *Engineering Applications of Artificial Intelligence* 12.4. 401–409. URL: http://www.sciencedirect.com/science/article/B6V2M-3YYT697-1/2/cb6830be50339d434887a6d9d8ab9200.

Forlizzi, Jodi, and Carl DiSalvo (2006). "Service robots in the domestic environment: a study of the roomba vacuum in the home". In: *HRI '06: Proceeding of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. Salt Lake City, Utah, USA: ACM Press. ISBN 1-59593-294-1. DOI: http://doi.acm.org/10.1145/1121241.1121286. 258–265.

Foundation, Blender (2007). *blender.org*. 2007. URL: http://www.blender.org/ (visited on 08/01/2007).

Freud, Sigmund (1953–1974 [1917] [1915]). "Mourning and Melancholia". In: *The Standard Edition of the Complete Psychological Works of Sigmund Freud*. Ed. by James Strachey. London, UK: The Hogarth Press. 239–258.

— (2003 [1919]). "The Uncanny". In: *The Uncanny*. Penguin Books. 121–161.

Gruau, Frederic, and Kameel Quatramaran (1996). Cellular encoding for interactive evolutionary robotics. Tech. rep. Amsterdam, The Netherlands, The Netherlands.

Haraway, Donna J. (1991). "A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century". In: *Simians, Cyborgs, and Women: The Reinvention of Nature*. Routledge.

Hashimi, Sama'a Al, and Gordon Davies (2006). "Vocal telekinesis: physical control of inanimate objects with minimal paralinguistic voice input". In: *MULTIMEDIA '06: Proceedings of the 14th annual ACM international conference on Multimedia*. Santa Barbara, CA, USA: ACM Press. ISBN 1-59593-447-2. DOI: http://doi.acm.org/10.1145/1180639.1180822. 813–814.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2001). *The Elements of Statistical Learning*. Springer.

Hayles, N. Katherine (2001). "Desiring Agency: Limiting Metaphors and Enabling Constraints in Dawkins and Deleuze/Guattari". In: *SubStance* 30.1 & 2. 144–159.

Heider, Fritz, and Marianne Simmel (1944). "An Experimental Study of Apparent Behavior". In: *The American Journal of Psychology* 57.2. 243–259. ISSN 0002-9556. URL: http://links.jstor.org/sici?sici=0002-9556%28194404%2957%3A2%3C243%3AAESOAB%3E2.0.CO%3B2-B.

Hurst, Jacob, and Larry Bull (2006). "A Neural Learning Classifier System with Self-Adaptive Constructivism for Mobile Robot Control". In: *Artificial Life* 12.3. 353–380. DOI: 10.1162/artl.2006.12.3.353. URL: http://www.mitpressjournals.org/doi/abs/10.1162/artl.2006.12.3.353.

Ishii, K., T. van der Zant, V. Becanovic, and P. Ploger (2004). "Optimization of parameters of echo state network and its application to underwater robot". In: *Proceedings of the SICE 2004 Annual Conference*. Vol. 3. 2800–2805 vol. 3.

Jaeger, Herbert (2001*a*). Short Term Memory in Echo State Networks. GMD 152. Tech. rep. Bremen, Germany: German National Research Center for Information Technology.

— (2001*b*). The "Echo State" Approach to Analysing and Training Recurrent Neural Networks. GMD 148. Tech. rep. Bremen, Germany: German National Research Center for Information Technology.

Jakobi, Nick (1998). "Minimal Simulations For Evolutionary Robotics". PhD thesis. University of Sussex.

Jehan, Tristan (2005). "Creating Music by Listening". PhD thesis. Massachusetts Institute of Technology.

Johnson, Colin G. (1999). "Exploring the sound-space of synthesis algorithms using interactive genetic algorithms". In: *Proceedings of the AISB'99 Symposium on Musical Creativity*. 20–27.

Kidd, C. D., and C. Breazeal (2004). "Effect of a robot on user perceptions". In: *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on* 4. 3559–3564 vol.4.

Kidd, Cory D. (2003). "Sociable Robots: The Role of Presence and Task in Human-Robot Interaction". Master's thesis. Massachusetts Institute of Technology.

Kidd, Cory D., Will Taggart, and Sherry Turkle (2006). "A Sociable Robot to Encourage Social Interaction Among the Elderly". In: *International Conference on Robotics and Automation*.

Kim, Hee-Su, and Sung-Bae Cho (2000). "Application of interactive genetic algorithm to fashion design". In: *Engineering Applications of Artificial Intelligence* 13.6. 635–644. URL: http://www.sciencedirect.com/science/article/B6V2M-41TN609-3/2/f36f199cd3de20d072879d123e3a04ff.

Kim, Sujeong, Xinyu Zhang, and Young J. Kim (2006). "Haptic Puppetry for Interactive Games". In: *Proceedings of Edutainment 2006*. 1292–1302.

Kleist, Heinrich von (1972). "On the Marionette Theatre". In: *The Drama Review: TDR* 16.3. 22–26. ISSN 0012-5962. URL: http://links.jstor.org/sici?sici=0012-5962%28197209%2916%3A3%3C22%3AOTMT%3E2.0.CO%3B2-B.

Kornblut, Anne E. (2006). "The Elephant in the Room". In: *New York Times* (29 2006).

Kristeva, Julia (1994). *Strangers to Ourselves*. New York, NY, USA: Columbia University Press.

Lapointe, François-Joseph (2005). "Choreogenetics: the generation of choreographic variants through genetic mutations and selection". In: *GECCO '05: Proceedings of the 2005 workshops on Genetic and evolutionary computation*. Washington, D.C.: ACM Press. DOI: http://doi.acm.org/10.1145/1102256.1102338. 366–369.

Lapointe, François-Joseph, and Martine Époque (2005). "The dancing genome project: generation of a human-computer choreography using a genetic algorithm". In: *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*. Hilton, Singapore: ACM Press. ISBN 1-59593-044-2. DOI: http://doi.acm.org/10.1145/1101149.1101276. 555–558.

Latour, Bruno (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Clarendon Lectures in Management Studies. New York, NY, USA: Oxford University Press.

— (1993). *We Have Never Been Modern*. Harvard University Press.

Latour, Bruno, and Steve Woolgar (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press.

Law, John (2004). *After Method: Mess in social science research*. New York, NY, USA: Routledge.

Lee, Sau lai, Ivy Yee man Lau, S. Kiesler, and Chi-Yue Chiu (2005). "Human Mental Models of Humanoid Robots". In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. 2767–2772.

Levin, Golan, and Zachary Lieberman (2004). "In-Situ Speech Visualization in Real-Time Interactive Installation and Performance". In: *Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering*. ACM. Annecy, France.

Lewis, M. A., A. H. Fagg, and A. Solidum (1992). "Genetic programming approach to the construction of a neural network for control of a walking robot". In: *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*. 2618–2623 vol.3.

Lippit, Akira Mizuta (2000). *Electric Animal: Toward a Rhetoric of Wildlife*. Minneapolis, MN, USA: University of Minnesota Press.

Lipson, Hod (2006). "Evolutionary Design and Evolutionary Robotics". In: *Biomimetics*. Ed. by Yoseph Bar-Cohen. Boca Raton, FL, USA: CRC Press.

Lipson, Hod, and Jordan B. Pollack (2000). "Automatic design and manufacture of robotic lifeforms". In: *Nature* 406 (31 2000). 974–978.

Lund, Henrik Hautop (2003). "Co-evolving Control and Morphology with LEGO Robots". In: *Morpho-functional Machines: The New Species: Designing Embodied Intelligence*. Ed. by F. Hara and Rolf Pfeifer. Springer-Verlag.

Lund, Henrik Hautop, Orazio Miglino, Luigi Pagliarini, Aude Billard, and Auke Ijspeert (1998). "Evolutionary Robotics—A Children's Game". In: *Proceedings of the IEEE 5th International Conference on Evolutionary Computation*. IEEE Press.

Macinnes, I., and E. Di Paolo (2004). "Crawling Out of the Simulation: Evolving Real Robot Morphologies Using Cheap Reusable Modules". In: *Artificial Life IX: Proceedings of the Ninth Interational Conference on the Simulation and Synthesis of Life*. Cambridge, Massachusetts: MIT Press. 94–99. URL: `citeseer.ist.psu.edu/700266.html`.

McCormack, Jon (2001). "Eden: An Evolutionary Sonic Ecosystem". In: *Advances in Artificial Life : 6th European Conference, ECAL 2001, Prague, Czech Republic, September 10-14, 2001, Proceedings*. 133–. URL: `http://www.springerlink.com/content/09h2c5a724d74ea2`.

McCormack, Jon (2002). "Evolving for the Audience". In: *International Journal of Design Computing* 4. URL: http://faculty.arch.usyd.edu.au/kcdc/ijdc/vol04/papers/mcormackFrameset.htm.

— (2006). "New challenges for evolutionary music and art". In: *SIGEVOlution* 1.1. 5–11. ISSN 1931-8499. DOI: http://doi.acm.org/10.1145/1138470.1138472.

McDermott, James, Niall J. L. Griffith, and Michael O'Neill (2005). "Toward User-Directed Evolution of Sound Synthesis Parameters". In: *EvoWorkshops 2005*. Springer-Verlag. 517–526. URL: http://www.springerlink.com/content/ucf8564mt4nphyr3.

Michalowski, Marek P., Selma Sabanovic, and Hideki Kozima (2007). "A dancing robot for rhythmic social interaction". In: *HRI '07: Proceeding of the ACM/IEEE international conference on Human-robot interaction*. Arlington, Virginia, USA: ACM Press. ISBN 978-1-59593-617-2. DOI: http://doi.acm.org/10.1145/1228716.1228729. 89–96.

Michalowski, Marek P., Selma Sabanovic, and Philipp Michel (2006). "Roillo: Creating a Social Robot for Playrooms". In: *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006)*.

Mouffe, Chantal (2000). *The Democratic Paradox*. Verso.

Mrdeza, S., and I. S. Pandzic (2003). "Analysis of virtual reality contribution to treatment of patients with post-traumatic stress disorder". In: *Proceedings of the 7th International Conference on Telecommunications*. Vol. 2. 729–731 vol.2.

Murdoch, Michael J. (2006). "Nonverbal Vocal Interface". Master's thesis. Rochester Institute of Technology.

Mutlu, Bilge (2006). "An empirical framework for designing social products". In: *DIS '06: Proceedings of the 6th ACM conference on Designing Interactive systems*. University Park, PA, USA: ACM Press. ISBN 1-59593-367-0. DOI: http://doi.acm.org/10.1145/1142405.1142456. 341–342.

Mutlu, Bilge, Jodi Forlizzi, Illah Nourbakhsh, and Jessica Hodgins (2006). "The use of abstraction and motion in the design of social interfaces". In: *DIS '06: Proceedings of the 6th ACM conference on Designing Interactive systems*. University Park, PA, USA: ACM Press. ISBN 1-59593-367-0. DOI: http://doi.acm.org/10.1145/1142405.1142444. 251–260.

Nass, Clifford, and Youngme Moon (2000). "Machines and Mindlessness: Social Responses to Computers". In: *Journal of Social Issues* 56.1. 81–103. DOI: 10.1111/0022-4537.00153. URL: http://www.blackwell-synergy.com/doi/abs/10.1111/0022-4537.00153.

Nolfi, Stefano, and Dario Floreano (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. Cambridge, MA, USA: MIT Press.

Oliver, William David (1997). "The Singing Tree: A Novel Interactive Musical Experience". Master's thesis. Massachusetts Institute of Technology.

Oliver, William, John Yu, and Eric Metois (1997). "The Singing Tree: Design of An Interactive Musical Interface". In: *Proceedings of DIS '97*. 261–264.

Orr, Scott P., Zahava Solomon, Tuvia Peri, Roger K. Pitman, and Arieh Y. Shalev (1997). "Physiologic responses to loud tones in israeli veterans of the 1973 yom kippur war;". In: *Biological Psychiatry* 41.3. 319–326. URL: http://www.sciencedirect.com/science/article/B6T4S-3RH12C7-X/2/84e973defc576fbd55cdaf74148654b3.

Painter, T., and A. Spanias (2000). "Perceptual coding of digital audio". In: *Proceedings of the IEEE* 88.4. 451–515.

Pair, J., B. Allen, M. Dautricourt, A. Treskunov, M. Liewer, K. Graap, G. Reger, and A. Rizzo (2006). "A Virtual Reality Exposure Therapy Application for Iraq War Post Traumatic Stress Disorder". In: *Proceedings of the IEEE VR2006 Conference*. Vol. 64–71. Alexandria, Virginia.

Papanek, Victor (1985). *Design for the Real World: Human Ecology and Social Change*. Academy Chicago Publishers.

Peeters, Geoffroy (2004). A large set of audio features for sound description (similarity and classification) in the CUIDADO project. Tech. rep. IRCAM.

Penny, Simon (1998). "Embodied cultural agents and synthetic sociology (poster)". In: *AGENTS '98: Proceedings of the second international conference on Autonomous agents*. Minneapolis, Minnesota, United States: ACM Press. ISBN 0-89791-983-1. DOI: http://doi.acm.org/10.1145/280765.280903. 463–464.

— (1997). "Embodied cultural agents: at the intersection of robotics, cognitive science and interactive art". In: *Proceedings of the AAAI Socially Intelligent Agents Symposium*. AAAI. 103–105. URL: http://www.ace.uci.edu/penny/texts/Embodied_Cult_Agents.html.

Piaget, Jean (1951 [1929]). *The Child's Conception of the World*. Littlefield, Adams.

Pickering, Andrew (1993). "The Mangle of Practice: Agency and Emergence in the Sociology of Science". In: *The American Journal of Sociology* 99.3. 559–589. ISSN 0002-9602. URL: http://links.jstor.org/sici?sici=0002-9602%28199311%2999%3A3%3C559%3ATMOPAA%3E2.0.CO%3B2-6.

Plöger, Paul G., Adriana Arghir, Tobias Günther, and Ramin Hosseiny (2004). "Echo State Networks for Mobile Robot Modeling and Control". In: *RoboCup 2003: Robot Soccer World Cup VII*. 157–168. URL: http://www.springerlink.com/content/88yex45c3jcx81x6.

Pollack, Jordan B., Gregory S. Hornby, Hod Lipson, and Pablo Funes (2003). "Computer Creativity in the Automatic Design of Robots". In: *Leonardo* 36.2. 115–121. DOI: 10.1162/002409403321554170. URL: http://www.mitpressjournals.org/doi/abs/10.1162/002409403321554170.

Proschan, Frank (1983). "The semiotic study of puppets, masks, and performing objects". In: *Semiotica* 47.1/4. 3–44.

Reeder, Roberta (1989). "Puppets: Moving Sculpture". In: *The Journal of Decorative and Propaganda Arts* 11. 106–125. ISSN 0888-7314. URL: http://links.jstor.org/sici?sici=0888-7314%28198924%2911%3C106%3APMS%3E2.0.CO%3B2-R.

Rizzo, A., J. Pair, K. Graap, P. McNerney, B. Wiederhold, M. Wiederhold, and J. Spira (2006). "A Virtual Reality Exposure Therapy Application for Iraq War Military Personnel with Post Traumatic Stress Disorder: From Training to Toy to Treatment". In: *NATO Advanced Research Workshop on Novel Approaches to the Diagnosis and Treatment of Posttraumatic Stress Disorder*. IOS Press. 235–250.

Rizzo, A., J. Pair, P. McNerney, E. Eastlund, B. Manson, and J Buckwalter (2005). "Design and Development of a VR Therapy Application for Iraq War Veterans with PTSD". In: *Technology and Informatics*. Ed. by J.D. Wstwood et al. Amsterdam, NL: OIS Press. 407–413.

Rothbaum, Barbara Olasov, Larry Hodges, Renato Alarcon, David Ready, Fran Shahar, Ken Graap, Jarrel Pair, Philip Hebert, Dave Gotz, Brian Wills, and David Baltzell (1999). "Virtual Reality Exposure Therapy for PTSD Vietnam Veterans: A Case Study". In: *Journal of Traumatic Stress* V12.2. 263–271. URL: http://dx.doi.org/10.1023/A:1024772308758.

Russolo, Luigi (2004 [1913]). "The Art of Noises: Futurist Manifesto". In: *Audio Culture*. Ed. by Christoph Cox and Daniel Warner. New York, NY, USA: Continuum. 370–380.

Salmen, M., and P. G. Ploger (2005). "Echo State Networks used for Motor Control". In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. Barcelona, Spain. 1953–1958.

Scarry, Elaine (1985). *The Body in Pain: The Making and Unmaking of the World*. New York, NY, USA: Oxford University Press.

Schaefer, Margret (1978). "Psychoanalysis and the Marionette Theater: Interpretation Is Not Depreciation". In: *Critical Inquiry* 5.1. 177–188. ISSN 0093-1896. URL: http://links.jstor.org/sici?sici=0093-1896%28197823%295%3A1%3C177%3APATMTI%3E2.0.CO%3B2-1.

Scholl, B J, and P D Tremoulet (2000). "Perceptual causality and animacy". In: *Trends Cogn Sci* 4.8. 299–309. URL: http://www.hubmed.org/display.cgi?uids=10904254.

Schumann, P. (1991). "The radicality of the puppet theater". In: *TDR: The Drama Review* 35.4. 75–83. ISSN 10542043. URL: http://search.ebscohost.com/login.aspx?direct=true&db=lfh&AN=9202172627&site=ehost-live.

Sengers, Phoebe, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye (2005). "Reflective design". In: CC *'05: Proceedings of the 4th decennial conference on Critical* *computing*. Aarhus, Denmark: ACM Press. ISBN 1-59593-203-8. DOI: http://doi.acm.org/10.1145/1094562.1094569. 49–58.

Sengers, Phoebe, and Bill Gaver (2005). "Designing for Interpretation". In: *Proceedings of Human-Computer Interaction International*.

— (2006). "Staying open to interpretation: engaging multiple meanings in design and evaluation". In: *DIS '06: Proceedings of the 6th ACM conference on Designing Interactive systems*. University Park, PA, USA: ACM Press. ISBN 1-59593-367-0. DOI: http://doi.acm.org/10.1145/1142405.1142422. 99–108.

Sengers, Phoebe, Rainer Liesendahi, Werner Magar, Christoph Seibert, Boris Müller, Thorston Joachims, Weidong Geng, Pia Maartensson, and Kristina Höök (2002). "The enigmatics of affect". In: *DIS '02: Proceedings of the conference on Designing interactive systems*. London, England: ACM Press. ISBN 1-58113-515-7. DOI: http://doi.acm.org/10.1145/778712.778728. 87–98.

Sims, Karl (1991). "Artificial evolution for computer graphics". In: *SIGGRAPH '91: Proceedings of the 18th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press. ISBN 0-89791-436-8. DOI: http://doi.acm.org/10.1145/122718.122752. 319–328.

— (1994). "Evolving 3D morphology and behavior by competition". In: *Artif. Life* 1.4. 353–372. ISSN 1064-5462.

Smith, Russel (2007). *Open Dynamics Engine*. 2007. URL: http://ode.org (visited on 05/28/2007).

— (2006). *Open Dynamics Engine v0.5 User Guide*.

Squartini, Stefano, Stefania Cecchi, Michele Rossini, and Francesco Piazza (2007). "Echo State Networks for Real-Time Audio Applications". In: *Advances in Neural Networks –ISNN 2007*. 731–740. URL: http://dx.doi.org/10.1007/978-3-540-72395-0_90.

Stang, Pascal (2005). *Procyon AVRlib*. 2005. URL: http://hubbard.engr.scu.edu/avr/avrlib/ (visited on 09/2005/20).

Sussman, Mark (2001). "Performing the Intelligent Machine: Deception and Enchantment in the Life of the Automaton Chess Player". In: *Puppets, Masks, and Performing Objects*. Ed. by John Bell. TDR Books. MIT Press.

Taggart, Will, Sherry Turkle, and Cory D. Kidd (2005). "An Interactive Robot in a Nursing Home: Preliminary Remarks". In: *Proceedings of the CogSci Workshop on Android Science*. 56–61.

Takagi, H. (2001). "Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation". In: *Proceedings of the IEEE* 89.9. 1275–1296.

Thrift, Nigel (2004). "Electric animals". In: *Cultural Studies* 18.2 - 3. 461–482. URL: http://dx.doi.org/10.1080/0950238042000201617.

Turkle, Sherry (2006). "A Nascent Robotics Culture: New Complicities for Companionship". In: *American Association for Artificial Intelligence (AAAI)*.

— (1995). *Life on the Screen: Identity in the Age of the Internet*. New York, NY, USA: Touchstone.

— (2005). "Relational Artifacts/Children/Elders: The Complexities of CyberCompanions". In: *Proceedings of the CogSci Workshop on Android Science*. 56–61.

Wang, Jyhwen, Steve Liu, and Angie Price (2006). "The Dancing Marionette - An Interdisciplinary Capstone Design Experience for Engineering Technology and Computer Science Students". In: *Proceedings of ASEE 2006*.

Watson, R. A., S. G. Ficiei, and J. B. Pollack (1999). "Embodied evolution: embodying an evolutionary algorithm in a population of robots". In: *Proceedings of the 1999 Congress on Evolutionary Computation*. Vol. 1. –342 Vol. 1.

Wei, Sha Xin (2002). "Resistance Is Fertile: Gesture and Agency in the Field of Responsive Media". In: *Configurations* 10. 439–472.

Weiss, Allen S. (1995). *Phantasmic Radio*. Durham, NC, USA: Duke University Press.

Whitman, Brian A. (2005). "Learning the Meaning of Music". PhD thesis. Massachusetts Institute of Technology.

Winnicott, D. W. (2005 [1971]*a*). "The Location of Cultural Experience". In: *Playing and Reality*. New York, NY, USA: Routledge. 128–139.

— (1969). "The Use of an Object". In: *International Journal of Psychoanalysis* 50. 711–716.

— (2005 [1971]*b*). "Transitional Objects and Transitional Phenomena". In: *Playing and Reality*. New York, NY, USA: Routledge. 1–34.

Wright, Matthew, and Adrian Freed (1997). "Open Sound Control: A New Protocol for Communicating with Sound Synthesizers". In: *Proceedings of the International Computer Music Conference*.

Wright, Matthew, Adrian Freed, and Ali Momeni (2003). "Open Sound Control: State of the Art 2003". In: *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*. 153–159.

Yamane, K., J. K. Hodgins, and H. B. Brown (2003). "Controlling a marionette with human motion capture data". In: *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*. Vol. 3. ISBN 1050-4729. 3834–3841 vol.3.

Yanase, Toshihiko, and Hitoshi Iba (2006). "Evolutionary motion design for humanoid robots". In: *GECCO '06: Proceedings of the 8th annual conference on Genetic and evolutionary computation*. Seattle, Washington, USA: ACM Press. ISBN 1-59593-186-4. DOI: http://doi.acm.org/10.1145/1143997.1144291. 1825–1832.

Zwicker, E., and E. Terhardt (1980). "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency". In: *The Journal of the Acoustical*

Society of America 68.5. 1523–1525. DOI: 10.1121/1.385079. URL: http://link.aip.org/link/?JAS/68/1523/1.

Index

Colophon

This thesis was produced using LaTeXand the `memoir` style with many customizations. Additional packages include `bbding`, `biblatex`, `glossary`, and `microtype`. Body text, chapter titles, and section headings are set in 11pt Adobe Jenson Pro; figure and table captions in 11pt Cronos MM; and monospaced text in 11pt Bitstream Vera Sans Mono. The text in the figures is DINSchrift Regular. Where possible I have tried to use Creative Commons Licensed graphics.