UNIVERSITY OF CALIFORNIA,
IRVINE

Computational Generation of Dream-like Narrative: Reflections on The Uncanny Dream Machine

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Information and Computer Science

by

Sean E. Voisen

Thesis Committee:
Professor Simon Penny, Chair
Professor Beatriz da Costa
Professor Paul Dourish

2010

Dedication

To my parents and my beloved wife Sara

for their patience and unfaltering support

Table of Contents

List of Figures

Acknowledgments

Simon Penny – Your continuous stream of reading suggestions, philosophical guidance, and bottomless pots of tea helped make this work possible. Thank you for helping me hone my ideas while providing new and interesting directions to explore and follow. And thank you also for Wall-Eye, which improved my soldering and circuit board layout skills beyond anything I could have achieved on my own. I will always fondly remember our sailing adventures, and look forward to the opportunity to partake in more.

Beatriz da Costa – Your willingness to have me work on your many and varied art projects is what made my ACE experience all that I hoped it would be. I will look back on those many late nights spent debugging or wandering around my neighborhood searching for a GPS signal and laugh heartily. Thank you for your academic, financial, and spiritual support. I wouldn't have made it through ACE in such good shape without it.

Paul Dourish – Thank you for your flexibility amidst a very busy schedule. Your instruction and guidance means a lot to me.

Tom Jennings – I left ACE with more practical skills and obscure knowledge than I ever could have asked for because of you. I will miss our spontaneous conversations about obsolete Cold War technology, animal consciousness, nixie tube physics, and the like. Thanks for the free stuff.

Robert, Brett, David and Eric – Thank you for all your camaraderie and support during my time in the ACE program.

Sara – Your unwavering patience during my many days of grumpiness and stress-induced insanity during these past two years is more than I could have ever asked for. You truly are the love of my life.

Mom and Dad – Your love and emotional support carried me through it all. I wouldn't have made it without you.

Abstract of the Thesis

Computational Generation of Dream-like Narrative:

Reflections on The Uncanny Dream Machine

By

Sean E. Voisen

Master of Science in Information and Computer Science

University of California, Irvine, 2010

Professor Simon Penny, Chair

The practice of expressive artificial intelligence (AI) involves the application of traditional AI practices to cultural production. The Uncanny Dream Machine is a work of expressive AI that employs these practices in order to examine and reflect on the phenomenon of nighttime dreaming. Drawing on computational linguistics and narrative techniques, The Uncanny Dream Machine uses a corpus of human dream transcripts to generate short, dream-like narratives that are uncannily similar to those found in the corpus. In so doing, it provides necessary practical grounding for deep theoretical questions about the interplay between dreams, narrative, creativity and embodiment in computation and everyday human experience.

Chapter 1: Introduction

1.1 Overview

What would it mean for a computer to dream? It is both an absurd and alluring question. It is absurd, from the standpoint of computer engineering, in that the answer to the question – and even the question itself – has no practical value. For most engineers the computer represents nothing more than a machine, an assemblage of dumb electronic componentry the purpose of which is to efficiently manipulate data – certainly not to dream. It is alluring, from the standpoint of philosophy, in that it offers a unique jumping off point for deep and even precocious questions about the nature of consciousness and the future possibility of self-aware machines. Could a machine dream? Do, as Philip K. Dick [12] once asked, androids dream of electric sheep? And if so, would we ever even be able to tell?

This text concerns itself with both the absurd and the alluring. It concerns itself both with computer science and philosophy, but also, and more primarily, with their intersections in specific cultural practices in the arts. In order to explore these intersections, I introduce and discuss "The Uncanny Dream Machine" an initial foray into computationally-generated dreamlike narrative conceived as a work of expressive artificial intelligence (*expressive AI*). Expressive AI, a term originally coined by Michael Mateas [30], describes "a new interdiscipline of AI-based cultural production" in which its practitioners combine art practice and artificial intelligence research into a cohesive whole. Expressive AI finds its inspiration from Philip Agre's [1] earlier notion of *critical technical practice*, in which cultural critiques of specific technical disciplines are expressed through technological innovation. Under the umbrella of expressive AI, critiques of AI find expression not as words on a printed page, but instead as tangible and embodied artifacts of electronics and code. The Uncanny Dream Machine is in some ways this: a critique of AI made tangible.

Why expressive AI? Interdisciplinary philosophical and technical investigations into AI of the kind that The Uncanny Dream Machine represents do not fit nicely within the field's normal cognitive frameworks. First and foremost, for the computer scientist working under traditional artificial intelligence paradigms, a dreaming machine is a useless machine. Most traditional artificial intelligence research concerns itself with problem solving. Traditional symbolist AI, sometimes called good old fashioned artificial intelligence – GOFAI for short – works under the assumption that the most valuable intelligence is the ability to solve problems. Interactionist AI, which is often portrayed as opposed to symbolist AI, operates under a similar assumption, adding only a single caveat: that the most valuable intelligence is the ability to solve problems *within a particular environment or context*. A dreaming computer, though it may exhibit intelligent behavior, and though the construction of which may advance the state of knowledge within the field, solves no problems. In fact, it may actually introduce more problems than it solves! As such, under the strict rubrics of both symbolist and interactionist AI, the act of building dreaming machines does not necessarily constitute a worthwhile pursuit.

Expressive AI offers one avenue for AI artists to follow which allows them to avoid the particular limitations of these traditional paradigms. Expressive AI does not concern itself with problem solving. Instead, the focus lies in expressivity, authorship and communication. Expressive AI concerns itself with building cultural artifacts that communicate a specific constellation of ideas through seemingly intelligent behavior. As such, whereas symbolist AI and interactionist AI focus on building artifacts that *are* independently intelligent, expressive AI focuses on building artifacts that "*seem* intelligent." Mateas writes, "Expressive AI views a system as performance. Within a performative space the system expresses the author's ideas. The system is both a messenger for and a message from the author." [30]

The Uncanny Dream Machine is a performance. It is a digital storyteller whose messages are more than mere dreams; they are also statements about the value and necessity of narrative, the impossibility of computational creativity, and the utter mystery of the stories we tell ourselves every morning we wake up after a night of image-rich slumber.

Consider this text a three-part "interpretive guide" to The Uncanny Dream Machine, its statements, and the lessons learned in the machine's creation. In the first part I offer some brief theoretical background on both the theory of dreams as narratives and the necessity of narrative in human culture. Drawing on this background, I propose that dreams, once communicated, not only *are* narratives but that they may in fact also serve as the *ultimate predecessor* of all fictional narratives. I then argue that in order to dream (and thus in order to write creative fiction) a dreamer needs more than a sleeping brain – he, she or it needs a sleeping brain embedded within a sleeping body. In the second part I expand on these arguments, demonstrating that because computers do not possess bodies, and thus cannot dream, their capacity to generate even the most rudimentary story remains limited. In so doing, I look at the history of computational narrative, explicate the challenges of symbol grounding [21] and narrative intelligence in computers, and explore the application of Margaret Boden's theories of creativity [4] to computational narrative. Finally, I conclude with a detailed description of The Uncanny Dream Machine itself, showing that though the art of computational narrative faces many challenges under significant handicap, through the lens of expressive AI such obstacles actually serve as assets for artists working at the intersection of story and code.

1.2 The Uncanny Dream Machine: A Brief Introduction

The Uncanny Dream Machine provides the tangible foundation upon which I anchor all of the forthcoming arguments and theoretical investigations. It is the work of embodied practice in which I ground any and all theorizing. This is in keeping with the original spirit of critical technical practice. As Agre writes, "A critical technical practice will, 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." [1] The Uncanny Dream Machine is presently where my first foot happens to be.

Though the final chapter of this text discusses the technical design and operation of The Uncanny Dream Machine in explicit detail, I first present here a general description of the machine. In so doing, I hope to give a sense not only of what the machine *is,* but also – and more importantly – what it *is not.*

Figure 1.1 The Uncanny Dream Machine

Embodied as a 1940's-era wooden Philco radio, The Uncanny Dream Machine allows listeners to “tune in” to various dream-like “streams of consciousness” delineated by one of eight different emotional qualities: fear, joy, curiosity, embarrassment, anger, anxiety, worry, and frustration. Once a listener selects a particular emotion using a knob on the radio, the machine generates a dream narrative characterized by that emotion and reads it back using a synthesized voice. As long as the listener remains on the selected emotional “channel,” the system continues to generate and read dreams of that emotional quality. Like a normal radio, the machine never stops talking unless it is switched off, continually generating and repeating new dreams ad infinitum.

To get a sense of the kind of dream narratives that the machine is capable of producing, consider the following sample narrative characterized by the emotion *frustration*:

> The first thing I remember is being on a beach trying to find my wallet. For some reason, I can't remember where I left my wallet. Sara approaches and says, "The wallet was stolen by Sara." For some reason I get the idea that Sara is lying. Sara denies my accusations. All of a sudden, my wallet appears in front of me.

The machine relies on a large corpus of real, recorded human dream narratives in order to determine the locations, people, objects and concepts that appear in its dreams. The more frequently a person, place or object appears in the corpus, the more likely it will also appear in the machine's dreams. Generated dreams retain much of the character and content of those found in the original corpus, but also introduce new scenarios, ideas, concepts and themes. It is this facet of the machine's design that gives it its uncanny quality; the dreams it generates are often uncannily similar to, but also different from, the corpus' original dreams.

In the sample above, the machine has drawn upon a collection of my own dreams, fifty-five in total, recorded over a period of six months as I could remember them upon waking. Sara, my wife, appears frequently in my own dreams, as she does in the dream above, and in a high proportion of dreams generated by the machine. Similarly, having resided in Southern California for most of my life, many of my own dreams take place at the beach or near the ocean, and as such so do many of the dreams generated by the machine, also like the one above.

In this way, the design and function of The Uncanny Dream Machine asserts the possibly contentious notion that dreams – at least as we remember them – are nothing but stories we tell ourselves when we awake. I expand on this contentious assertion in the following chapter, but suffice it to say that, in a nutshell, The Uncanny Dream Machine is an artificial intelligence system that generates short stories of a very particular genre – the genre of dreams.

Reiterating the fact that The Uncanny Dream Machine represents a work of expressive AI, focusing on the *appearance* of intelligence rather than *actual* intelligence, it should not be construed as a computational simulator of the dreaming process. That is, The Uncanny Dream Machine is not a study in the neurobiology of REM sleep nor an accurate attempt to model the human brain during nighttime dreaming. So little is known about the biology of dreaming that to attempt to computationally simulate it would be a work of hubris. Furthermore, as I expound upon later, computers are nothing but complicated manipulators of abstract symbolic tokens – they are not brains, they do not have bodies, and they do not have minds. As such, they cannot be expected to replicate the dreaming process in any way close to that of a human.

Nevertheless, from the outsider's point-of-view, The Uncanny Dream Machine truly is a dreaming machine. At least it *seems* that way. Given any particular transcript of output, it may be difficult (and sometimes maybe even impossible) to determine whether the transcript is the result of a real human's nighttime dive into the nether-regions of the subconscious unknown, or simply the result of a clever kind of computational prestidigitation. If all we have is a transcript, how do we judge?

1.3 The Expressive Turing Test

The challenge of differentiating between clever computational prestidigitation and real human creativity is analogous to that of differentiating between clever computational intelligence and real human intelligence. The Turing Test, originally proposed by Alan Turing in his seminal 1950 paper, "Computing Machinery and Intelligence," [54] offers one possible method for tackling this challenge. The test attempts to gauge an AI system's intelligence through careful human interlocution over a non-visual communications medium such as a telephone or written correspondence. During the Turing Test, a human interlocutor engages in conversation with a hidden and unknown interviewee using the chosen medium. The interlocutor's task is to determine whether said interviewee is a computer or a real human being. If after some extended period of in-depth questioning, the human interlocutor cannot determine if the responses to his or her questions have originated from a computer or a real human being then the computer is said to have passed the test. The computer is said to have real human-like intelligence.

Challenges to the Turing Test abound. Some, such as Searle's famous "Chinese Room Argument" [49] are nearly as well-known as the Turing Test itself. In the Chinese Room Argument, Searle questions the validity of the Turing Test with a simple thought experiment: Suppose that the test is taken in Chinese, and suppose that Searle (who knows no Chinese) assumes the role of the computer during the ensuing interlocution. Suppose also that a native Chinese speaker has secretly provided Searle with a large rulebook. This rulebook tells Searle how to take the incoming Chinese symbols, manipulate them based on their shape, and then output a new string of symbols such that it appears he understands the answers to the incoming questions. Would we say, then, that Searle must understand Chinese? Many agree with Searle that the answer to this question is an outright “no.”

Most arguments against the Turing Test, such as the Chinese Room Argument, claim that the test thus does not provide any means to gauge true intelligence, only the *appearance* of intelligence. Searle does not speak or understand a word of Chinese, but he can make it appear that he does. As such, because the Turing Test relies solely on a system's outward behavior, clever prestidigitation (like Searle and his Chinese rulebook) remains indistinguishable from real understanding.

Despite its flaws, one could also imagine a kind of Turing test for machine-generated dreams. Under such a test, the unlabeled transcripts of a real human dream and a computer-generated dream are presented to human judge. If the human judge cannot determine with better than a fifty percent chance which transcript was generated by a computer and which was generated by a human being, would we say that the computer is dreaming? In all likelihood, again the answer is probably “no.”

I revisit this line of questioning in Chapter 3, but at present consider the following: Expressive AI completely side-steps the problem that the Turing Test tries to solve. To reiterate, the focus of expressive AI remains on authorship, expressiveness through behavior, and the *appearance* of creativity, or the *appearance* of intelligence. The artist working with AI as an expressive medium does not care whether his or her creation is *actually* creative or *actually* intelligent, an attitude that renders the Chinese Room Argument and its kin moot. What remains is a new kind of Turing Test, identical in character, but of different intention: to determine the effectiveness of a system's behavior as an expressive system. I call this test the "Expressive Turing Test," and refer to it in the coming chapters.

1.4 Dreams: Syntax, Semantics and Story

Thus far I have shown that though The Uncanny Dream Machine may appear to dream, it does not engage in or simulate the real dreaming process. I must also stress that it does not generate psychologically meaningful dreams – not intentionally anyway. My concern in designing the machine was not to create a system that generated meaningful dreams in the Freudian sense, but to create a system that generated believable dreamlike *stories*. Believable dream stories require meaningfulness, but of an entirely different sort than that for which Freud devoted much of his career. They must be coherent. They must use understandable and grammatically correct language. They must describe concepts and activities that, when put together, make some semblance of sense, bizarre though they may be in certain juxtapositions. They need not carry hidden symbolism related to subconscious Oedipus complexes, suppressed sexual yearnings or other curious psychological disorders.

It should thus come as no surprise that this text does not concern itself with Freudian, Jungian or any other kind of psychoanalytic dream interpretation. I have no grudge against psychoanalysis or dream interpretation. Instead, I find them irrelevant to the discussion. They investigate a level of semantics several steps beyond that which I discuss here. Instead, when I discuss the "meaning" in a dream transcript, what I mean is something far more pedestrian and mundane: the semantics of a particular word or sentence. I concern myself here with dreams only at the most superficial level – in their syntax and meaningful coherence – and to legitimize them as a worthy study in the foundations of narrative. During this process of legitimization that I hope to show two things: First, why I have chosen dreams as a valid genre for computational narrative. And second, how through the lens of expressive AI the seemingly absurd futility of computationally generating narrative becomes less absurd.

The beauty of expressive AI, as with any other form of art, is that any additional layers of meaning beyond those I discuss here can be added and removed at will by the viewer. The machine itself has no clue about the meaning of its incessant babbling, psychoanalytically or otherwise. Computers have enough challenges understanding the meaning behind a single word, let alone the rich metaphor and symbolism that Freud might unveil in a single patient's dream. In fact, as I will soon argue, computers can't understand anything on any level of semantics. Meaning is meaningless to computers, and so The Uncanny Dream Machine, being at heart a computer, knows not what it says. It is precisely this which makes The Uncanny Dream Machine a very peculiar kind of storyteller. Like the specter of a disembodied schizophrenic rambling away at the heavens, its speeches find both humor and profundity only in the minds of the occasional passerby.

Chapter 2: Dreaming, Storytelling and Embodiment

2.1 Dreams as Involuntary Fiction

Storytelling is a fundamental aspect of the human experience. Stories infuse our lives from an early age, and soon after we learn to communicate we immerse ourselves in their riches. Fairy tales, picture books, cartoons, parental anecdotes, and even simple stories made up and told on the spot all constitute vivid and common childhood memories. Narrative is the way in which we make sense of the world; it is the tool we use to understand ourselves, understand our humanity, and communicate our findings to future generations [40]. It is how we create and sustain meaning, and we cannot live without it. In their paper on narrative intelligence, Mateas and Sengers [31] call narrative a "fundamental organizing principle of human existence," and citing Bruner, "the way in which we make sense of the world." As Reynolds Price writes in his book, *A Palpable God* [46]:

> A need to tell and hear stories is essential to the species Homo sapiens – second in necessity apparently after nourishment and before love and shelter. Millions survive without love or home, almost none in silence; the opposite of silence leads quickly to narrative, and the sound of story is the dominant sound of our lives, from the small accounts of our day's events to the vast incommunicable constructs of psychopaths.

Given the fundamental connection between narrative and our waking experience, it is little wonder that narrative plays an equally important role in our sleeping experience. That we directly experience our nighttime dreams as stories is debatable, but few would argue that we do not communicate our dreamtime experiences in the same way we communicate our waking experiences – as narrative. To do otherwise would seem absurd. Consider the following excerpt of a dream transcript – a personal dream of Freud's – taken from his own *The Interpretation of Dreams* [16]:

> A great hall – a number of guests, whom we are receiving – among them Irma, whom I immediately take aside, as though to answer her letter, and to reproach her for not yet accepting the "solution." I say to her: "If you still have pains, it is really only your own fault." She answers: "If you only knew what pains I have now in the throat, stomach, and abdomen – I am choked by them." I am startled, and look at her. She looks pale and puffy. I think that after all I must be overlooking some organic affection …

Consider also the following excerpt of a dream report by a six-year-old boy, taken from a case file of Jean Piaget's [44] during his early investigations into childhood development:

> In the basin I saw a bean that was so big that it quite filled it. It got bigger and bigger all the time. I was standing by the door. I was frightened. I wanted to scream and run away, but I couldn't. I got more and more frightened, and it went on until I woke up.

Or this dream report by a young girl, also from Piaget's [44] files:

> I dreamt that Mr. M fired a gun at a man who was high up in the air. The man was very ill and was going to die so he killed him. He was very small, and then when he fell he was big: he got bigger and bigger; he had a fat tummy like you; he was just like you!

We describe our dreams as anecdotes of personal experience because that is precisely what they are: personal experiences we generate for ourselves. Some say that in dreaming we can see our own thinking as it happens. Think of a lion and it magically appears. Think of a loved one and you may soon find yourself by his or her side. Think of an embarrassing situation and so it shall be made manifest, much to your chagrin. In his book, *Dreaming and Storytelling*, Bert States notes however, that “this is a somewhat slippery notion because you can also think thoughts about what you are seeing in a dream (which is also what you are thinking).” [52] This level of “meta-thinking” differentiates dream thought from waking thought in a very specific way. Whereas in waking life we concern ourselves *only* with the meta level of thought – that is, with thinking about what we see and experience – in dream thought we may simultaneously conjure up entire worlds *and* reflect upon them. In this way, dreamtime thought can be seen as analogous to the type of “meta-thinking” novelists engage in when simultaneously crafting entire fictional worlds and the thoughts and feelings of all the inhabitants therein.

States calls dreaming the “ur-form of all fiction.” [52] A form of uninhibited, self-directed storytelling, it precedes all written or oral fictions in both “the personal life of the individual” and “historical life of the species.” He notes that, after all, even animals dream. Some biological theories of dreaming claim that the phenomenon may have evolved as a kind of threat simulation mechanism [47]. These theories suggest that dreaming acts as a way of rehearsing responses to potentially life-threatening situations on the off-chance that they should arise in real-life. States writes, “this means, in a loose sense, that animals tell themselves stories, and the stories may be well about survival routines.” [52]

What is required to be able to dream? What does it require to conjure up the worlds, the vivid and rich imagery of the dreaming experience? Does it require conscious self-awareness? Does it require language? If animals really do dream, then the likely answer to at least the latter question is “no.” Language may be required to communicate the experience of a dream – just as a novelist uses the written word to communicate the experience of his or her own imagination – but language is perhaps not required to have the experience itself. Instead, I argue that the most fundamental component necessary for dreaming is, quite simply, a body.

2.2 The Umwelt and the Brain in a Vat

The idea of the “brain in the vat” [53] – the notion that one can somehow partition off the mind from the body – has pervaded Western philosophical discourse since Descartes famously pronounced, “cogito ergo sum” in 1637. The brain in a vat theory of cognition – more formally known as *cognitivism* – asserts the brain's independence as a thinking entity, viewing the body as nothing more than an input/output and locomotion mechanism for the brain. Cognitivism relies on the principle of *explanatory disembodiment*, “which states that the details of the physical embodiment of an agent can safely be ignored for the purposes of cognitive explanation.” [27] Under cognitivism, cognition becomes an act that exists purely within the realm of the abstract and mental; thinking becomes something that can only be explained and understood by examining the activities of the non-physical *mind*.

By embracing the Cartesian mind-body split, cognitivism has spawned a significant philosophical challenge for its adherents. The challenge, referred to as the “mind-body problem,” [19] lies in explaining both the origin of the non-physical mind and in explaining how such a non-physical entity can communicate with the otherwise physical bodily corpus. No trivial matter, the mind-body problem has yet to be solved.

Nevertheless, cognitivism and all the significant philosophical or scientific baggage it either spawned or reinforced – the mind-body problem, solipsism, the acceptance of the brain as the sole information processing center of the body, and the notion of thought as abstract symbol manipulation – has infused itself so subtly in Western culture that it sat largely unquestioned for hundreds of years. Only in the past century or so, particularly since the advent of the computer and recent failure of symbolist AI (which has foundations in cognitivism) to produce anything resembling human intelligence, has Cartesian dualist philosophy faced significant questioning. Its most significant opponent, the philosophy of *embodied cognition*, with its roots in Merleau-Ponty's phenomenology of embodiment, argues that rather than lying in abstract isolation away from the body, thought processes are actually shaped by the body and its interactions with the world. Put more succinctly, embodied cognition claims that, contrary to cognitivist beliefs, the mind is not just what the brain thinks. Instead, it purports that the world shapes the body, which in turn shapes the mind. This notion, that world shapes body shapes mind, finds some of its earliest and most significant support in the work of biologist Jakob von Uexküll [57].

Von Uexküll's “A Stroll Through the Worlds of Animals and Men,” now a foundational treatise in the nascent field of biosemiotics, explores the biological world from a subjective perspective. In it, Von Uexüll argues that proponents of “mechanistic theories” of biology “have overlooked the most important thing, the *subject,*” or the “operator,” of a biological body's “perceptual” and “effector” tools. Not to be confused with a kind of neo-Cartesianism, Von Uexküll does not see the subject or operator of a body as separate from that body, but rather "built into" its very organs. He describes a unified organism constrained by the intersection of the world it can perceive and the world it can change or effect. This intersection Von Uexküll calls the *umwelt,* now defined as the self-centered world that is unique to a particular organism and imperceptible by any organism outside the same species. In other words, the umwelt is a species-specific universe.

To give the reader a better picture of what he means by *umwelt*, Von Uexküll describes the umwelt of a common tick. The tick, blind and deaf though it may be, can nevertheless determine the presence of any warm-blooded mammal nearby merely through the scent of butyric acid, which "emanates through the skin glands of all mammals." [57] Using its keen and focused sense of smell as the only cue as to the location of its next victim, the tick will climb to the top of the nearest plant or blade of grass and lie in wait, sometimes for years, for its next – and probably last – meal. Obviously, ticks lives a vastly different subjective experience from our own. Just as we humans cannot detect the smell of butyric acid from our nearest neighbor (except in large quantities, in which it presents itself as common "body odor"), so the tick cannot detect the blue of the sky or the chirping of the nearest bird. Though we may share the same planet, humans and ticks live in entirely different universes, or more accurately, vastly different umwelts.

The concept of umwelt extends beyond biology, however. Our umwelt defines not only our experience of the external world, but also our inner imaginative experience. This means that we are capable of thinking about and imagining only that which is conceivable within our umwelt. For instance, the Papilio butterfly, having five different types of color receptors in its eyes, possesses pentachromatic vision capable of detecting colors in ultraviolet wavelengths [26]. How could we, as humans, with no experience of pentachromatic vision (we are only trichromatic), even imagine what such ultraviolet colors might look like? Try as we might, we cannot. Similarly, how could a tick “imagine” the color blue when it does not even have the experience of sight, let alone color? In this way the world, through evolutionary coupling and through our individual interactions with it, shapes our bodies which, in turn, shape and define our own thinking.

Many years after Von Uexküll, Francisco Varela and his colleagues used similar observations to develop the theory of *enactivism* [56]. Enactivism offers a more formal approach to the mind-body-world condition of cognition described above, rejecting outright the explanatory disembodiment of cognitivism. The theory proposes that cognition is dynamic, emerging “from the coupled interactions of the brain, body, and environment,” [10]. Hence, it cannot be understood by examining mental processes alone. Critical to the theory of enactivism is the notion of *structural coupling*, in which the body of an entity determines its form of engagement with its environment or other entities within that environment. According to Maturana and Varela, structural coupling is not a statically defined phenomenon, but instead evolves over “recurrent interactions” that lead to “the structural congruence between two (or more) systems.” [32]

This “structural congruence” is strikingly similar to Von Uexküll's observations about an organism's interactions with its umwelt via its “perceptual” and “effector” tools. What differentiates enactivism from Von Uexküll's umwelt theory is its stronger focus on how environment and body affect cognition and mind. Von Uexküll, as a biologist, is primarily concerned with the body's impact on an organism's subjective experience, and how that subjective experience allows an organism to detect or ignore environmental cues. Thus, while Von Uexküll does not reject cognitivism outright, his observations help plant the seeds that lead to its rejection.

An intellectual successor to umwelt theory, enactivism broadens and enhances it, stressing how an organism's umwelt not only constrains its subjective experience of the world, but also constrains and defines its thinking about that subjective experience.

2.3 The Umwelt and the Dreaming Body

Returning to my previous argument – that the act of dreaming requires a body – I propose this question: If our umwelt constrains and defines our thinking, then would it not also constrain and define our dreams? Does our umwelt affect more than just our subjective waking experience? I argue that the answer to this question is a profound "yes." Think back on your own dreams. Have you ever dreamt an experience that could not be explained in the context of your waking experience? That is, have you ever dreamt of having a new, sixth sense, or of occupying more than three dimensions, or of shrinking to the size of an ant, or – in the famous fictional case of Kafka's Gregor Samsa – of transforming into a giant, hissing cockroach? The bizarreness of certain dream experiences may lead us to believe that we truly can break out of the confines of our waking umwelt. Certainly the preceding examples, bizarre though they are, seem plausible. I suggest, however, that this bizarreness is misleading. However bizarre a dream experience may be, it can still only be understood within the context of one's waking life. I may dream, for example, that like Gregor Samsa I suddenly find myself transformed into a giant cockroach. What does this mean, exactly? Consider Kafka's own account of the experience [25]:

> He lay on his back, which was as hard as an armor plate, and raising his head a little, he could see the arch of his great, brown belly, divided by bowed corrugations. The bedcover was slipping helplessly off the summit of the curve, and Gregor's legs, pitiably thin compared with their former size, fluttered helplessly before his eyes.
>
> "What has happened?" he thought.

Certainly vivid, the trouble with Kafka's description lies in the fact that it retains a human's perspective. A real cockroach would certainly never think, "What has happened?" as it likely possesses neither the language nor mental faculties to do so. Considering the fact that Kafka wrote *The Metamorphosis* for a human audience, we can excuse his philosophical transgressions, as a novel solely composed of a cockroach's hissing and clicking noises would probably not make the bestseller list. Nevertheless, his approach to *The Metamorphosis* – indeed his entire premise – is undeniably Cartesian in nature. Here we find a mind partitioned from its body; the body transformed but the mind unaffected, as if it were somehow extracted and placed in a new, unwitting entomological host. The result is not a cockroach's experience of the world, but a human's experience of a cockroach's experience, itself a kind of *first-order strange loop*, to borrow a phrase from Douglas Hofstadter [23].

What is it really like to be a cockroach? We will likely never know. Thomas Nagel calls the "subjective character of experience" of a particular organism the "what it is like to be” that organism, pointing out that we can never know what it is like to be anything other than our own species [37]. By exploring “what it is like to be a bat,” Nagel discusses, for instance, the subjective experience of possessing sonar-based perceptual capabilities. He writes, “... bat sonar, though clearly a form of perception, is not similar in its operation to any sense that we possess, and there is no reason to suppose that it is subjectively like anything we can experience or imagine.” [37] If this sounds similar to the Papilio butterfly vision example cited previously, that is because it is.

Our subjective experience finds its limits in our body's perceptual and effectual capabilities. This is what Von Ueküll was pointing to when coining the term *umwelt,* and what Nagel attempts to demonstrate when arguing that we can never know what it is like to – objectively – be like a bat. While we can certainly imagine what it is like to inhabit another organism's umwelt, we can do so only within the context of our own. While we can certainly dream about transforming into a cockroach, our experience of being a cockroach remains largely human.

This results in the following two propositions: First, the umwelt of the dreaming brain is a subset of its larger, waking umwelt. Second, there is such a thing as an umwelt of the dreaming brain. While this should seem obvious – for we perceive and affect our dreams in a way very similar to our waking life – neurological studies of rapid eye movement (REM) sleep also lend support to the idea. REM sleep is the stage of sleep most commonly associated with the dreaming process. Studies of REM sleep have found that during this stage of the sleep cycle, the activity of the brain's neurons closely resembles that of the waking state [28]. In “Of Dreaming and Wakefulness” Llinás and

Paré write, "… paradoxical (REM) sleep and wakefulness are … almost identical intrinsic functional states in which subjective awareness is generated." Later they add, "… REM sleep can be considered as a modified attentive state in which attention is turned away from sensory input, toward memories … Let us formally propose then that wakefulness is nothing other than a dreamlike state modulated by the constraints produced by specific sensory inputs." Thus, while we are awake our umwelt is defined and constrained by our body's sensory inputs, but when we are asleep it is instead defined and constrained by our *memories* of our body's sensory inputs.

A "brain in a vat" has no umwelt. Borrowing the terminology of Nagel, there is no such thing as a what-it-is-like-to-be a brain in a vat. Without a body, it possesses neither the perceptual nor effectual tools with which to interact with and manipulate its environment. Without an umwelt, it cannot possibly have any kind of subjective experience, or if it does, such a subjective experience must consist of absolute nothingness. And without a subjective experience, a brain in a vat cannot be expected to dream. After all, the dream umwelt is a subset of the waking umwelt, but the subset of nothing is nothing. To quote a famous Zen koan: "A monk asked Jōshū, 'Has a dog the Buddha Nature?' Jōshū answered only, 'Mu.'" *Mu*, in Zen, literally means nothingness. Such is the subjective experience of the Cartesian mind separated from its body: nothingness, emptiness, a great and expansive void. And so, old and wise a Zen monk he was, we can only assume that if Jōshū today were asked, "Does a brain in a vat dream of anything?" his answer would be again only a terse, "Mu."

Chapter 3: Disembodied Zombie Storytellers

3.1 Computers are Not Like Brains in Vats

Though the analogy now seems ripe for use, computers are not quite like brains in vats. If they were to somehow be manufactured in such a way that they could be implanted in a person's skull, they likely could not serve as a viable substitute for biological grey matter. Similarly, brains in vats are not quite like computers. To reject cognitivism and embrace embodied cognition requires this assertion. Computers are fed information as symbolic tokens, themselves processed and encoded by an outside agent – typically a human – only to be regurgitated in modified form sometime later. Brains do not operate in this manner. There is no outside agent predigesting the information that the brain then processes. As Searle [50] argues, "the brain does not do information processing." For instance, the photo receptors on the back of a human retina do not send signals to the brain as symbolic tokens. Instead, they are sent as "electro-chemical reactions" that result in a near-simultaneous "conscious visual experience," itself a subjective, non-measurable phenomenon "as specific and concrete as a hurricane or the digestion of a meal."

This potential for phenomenological experience is the most substantive differentiator between brains and vats and computers. As real world implementations of Turing machines – devices capable of simulating any computable algorithm – computers can theoretically *simulate* the behavior of a brain in a vat. Researchers in artificial neural networks have sought to make this a reality for decades. Artificial neural networks are abstract computational models of biological neural networks, abbreviated simulations of the densely interconnected groups of neurons that constitute the nervous systems of biological organisms. Though computers do not possess neurons themselves, artificial neural networks make use of mathematical approximations of neural behavior expressed as computational algorithms to create “brain-like” behavior. To date, however, no artificial network has scaled to such complexity that it could accurately simulate the behavior of even a small mammalian brain, let alone a human brain. Nevertheless, ambitious projects such as the Blue Brain Project [29] hope to achieve this feat within a decade.

Yet, even if such projects do manage to create accurate simulations of the human brain, they still remain only simulations, and disembodied ones at that. The difference between the simulation of a brain, and the phenomenological subjective experience of an embodied real brain (mind), is perhaps best explained by examining the notion of the philosopher's zombie.

3.2 Computers are Like Disembodied Zombies

In invoking the term “zombie,” I am not attempting to conjure a Hollywood-style phantasmagoria of the waking dead. The idea behind the philosopher's zombie is foundationally similar, but substantially different in implementation and appearance. To better explain what is meant by the term *zombie* in the philosophical sense, Searle proposes a simple gedanken-experiment [51][6]: Suppose that due to cancer or some other illness your brain begins to deteriorate and that, desperate to save your life, doctors begin to replace each part of your diseased brain, piecemeal, with silicon-chip workalikes until only silicon remains. Searle describes three possible outcomes of such an experiment: the “smooth-as silk variation,” the “zombie variation,” and the “curare variation.” The “zombie variation” proceeds as follows:

> As the silicon is progressively implanted into your dwindling brain, you find that the area of your conscious experience is shrinking, but that this shows no effect on your external behavior. You find, to your total amazement, that you are indeed losing control of your external behavior … [You have become blind, but] you hear your voice saying in a way that is completely out of your control, 'I see a red object in front of me.' … We imagine that your conscious experience slowly shrinks to nothing, while your externally observable behavior remains the same.

Harnad [22] offers a more succinct definition of what it means to be a zombie: “If there's something home in there, something hurting when pinched, then that's a mind [and not a zombie].”

Computers are disembodied zombies. They are disembodied, according to Von Uekϋll's terminology, in that they do not possess the perceptual nor effectual bodily apparatus with which to sense or interact with the world. Or, according to the enactivist definition, in that they are not self-regulating systems comprising internal, homeodynamic processes, capable of sensorimotor coupling with the outside world [53].

They are zombies, according to Searle's gedanken-experiment, in that computers of sufficient sophistication can *act* as if “there's something home in there.” They can *simulate* self-conscious, self-aware behavior. But in the end, their “'inner lives' are no different than those of common, ordinary rocks.” [6] In short, just as there is no such thing as what-it-is-like-to-be a brain in a vat, so too is there no such thing as what-it-is-like-to-be a computer. This brains in vats and computers share in common: the lack of a subjective, conscious experience.

3.3 The ELIZA Effect

The computational simulation of conscious behavior sometimes renders itself so believable that it becomes easy to mistake such "zombie consciousness" for the real thing. Joseph Weizenbaum's [59] early experiment in computer-human dialogue, ELIZA, is perhaps the best-known example of such a phenomenon. An artificial intelligence program created partially as an experiment in natural language processing, and partially as a parody of Rogerian psychoanalysis, Weizenbaum published first published ELIZA in 1966. Consider the following transcript of a dialogue between ELIZA and a young woman, taken from Weizenbaum's *Computer Power and Human Reason*:

> *Woman:* Men are all alike.
> *ELIZA:* IN WHAT WAY
> *Woman:* They're always bugging us about something or other.
> *ELIZA:* CAN YOU THINK OF A SPECIFIC EXAMPLE
> *Woman:* Well, my boyfriend made me come here.
> *ELIZA:* YOUR BOYFRIEND MADE YOU COME HERE
> *Woman:* He says I'm depressed much of the time.
> *ELIZA:* I AM SORRY TO HEAR YOU ARE DEPRESSED
> *Woman:* It's true. I am unhappy.
> *ELIZA:* DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY
> *Woman:* I need some help, that much seems certain.

Shortly after its publication, ELIZA gained fame throughout artificial intelligence circles and popular culture as a beacon of hope and promise for the future of human-machine relationships. Though ELIZA's linguistic capabilities were limited at best, constrained to small sub-domains of interaction, the publicity surrounding ELIZA often ignored this fact. Some computer scientists believed that Weizenbaum had stumbled upon a general purpose solution to the problem of natural language processing.

Prominent psychiatrists lauded Weizenbaum's efforts at automating their practice. The renowned astrophysicist, Carl Sagan, praised the invention: “No such computer program is adequate for psychiatric use today, but the same can be remarked about some human psychotherapists. In a period when more and more people in our society seem to be in need of psychiatric counseling, and when time sharing of computers is widespread, I can imagine the development of a network of computer psychotherapeutic terminals, something like arrays of large telephone booths, in which, for a few dollars a session, we would be able to talk with an attentive, tested, and largely non-directive psychotherapist.” [59]

Weizenbaum penned *Computer Power and Human Reason* in hopes of tempering the public's almost unbridled enthusiasm towards his creation and its perceived implications for the future of computing. In it, Weizenbaum warns readers not to confound the simulation of human intelligence and emotion with the real things. He writes [59]:

> I was startled to see how quickly and how very deeply people conversing with [ELIZA] became emotionally involved with the computer and how unequivocally they anthropomorphized it ... What I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people. This insight led me to attach new importance to questions of the relationship between the individual and the computer, and hence to resolve to think about them.

So famous was ELIZA's propensity to this “induce powerful delusional thinking” that it lent the phenomenon its name. The human tendency to unintentionally anthropomorphize seemingly conscious behavior in computational systems is now called the "ELIZA effect."

3.4 Expressive AI and The ELIZA Effect

The ELIZA effect is a trap that can catch even the most rigorous thinkers. Critics of AI cite the ELIZA effect as a primary reason why AI researchers both overstate the capabilities and underestimate the shortcomings of their systems [14]. Consider even Alan Turing, who in the light of Searle's Chinese Room Argument, fell prey to a variation of the ELIZA effect when he proposed that the behavior of an intelligent system could serve as a viable measurement of its real, intrinsic intelligence.

Modern computer scientists may consider themselves immune to the kind of naivete exhibited by a young Turing during the earliest days of digital computing. But studies show that human behavior towards computers often inconsistently reflects an individual's espoused beliefs about the machines [39]. Such studies have found that even individuals who claim to know much about the inner-workings of computers nevertheless still tend to anthropomorphize and behave socially towards them. The reasons for this remain unclear. Recent research on speech-enabled and voice-activated computers suggests that humans unconsciously employ the same set of rules, cues and heuristics to determine their behavior towards "human-like" machines that they employ to determine their behavior towards other humans [38]. Drawing on evolutionary psychology, the suggestion is that humans will unconsciously use the most familiar mental framework (human sociality) to make predictions about foreign, nonhuman agent behavior. And that they do so more frequently the more human the behavior appears to be.

Much of the power of the ELIZA effect may also derive from the way system authors and observers ascribe intentional vocabulary to the workings of the systems. Zhu and Harrell [61] note that works of AI sometimes "... exhibit complex behaviors usually seen as the territory of intentional human phenomena, such as planning, learning, creating and carrying." The sheer act of stating that a machine "dreams" or "plans" or "creates" leads us down a narrational path that is both necessary and unavoidable, but also fraught with potential for unintentional deception.

From the standpoint of traditional artificial intelligence research then, the ELIZA effect poses a problem. If one's goal is to produce systems that actually *are* intelligent or actually *are* creative, and the only means of measuring intelligence or creativity is through behavior, how does one engage in objective measurement without the ELIZA effect's unconscious, subjective bias?

Fortunately, in the case of expressive AI, the dilemma ceases to exist. When one focuses on creating machines that are *seemingly* intelligent or *seemingly* creative, the ELIZA effect transforms from obstacle to asset, and one can side-step problem of objective assessment. Indeed, the goal of many works in expressive AI is often to induce the ELIZA effect to the greatest degree possible, to draw the viewer or participant into dramatic and fictional worlds rendered by clever code and behavior. Oftentimes, the greater the delusion the better the experience, as long as the viewer or participant willingly and knowingly participates. The fact that the machine inducing the delusion is actually a "disembodied zombie," unable to comprehend its own fabrications, is irrelevant. The behavior of the machine is nothing more than a medium, like text or video, a mechanism for delivering a message. In this way, creators of works in expressive AI walk in similar spheres as authors, playwrights, screenwriters and motion picture directors. Janet Murray, in *Hamlet on the Holodeck* for instance, likens Weizenbaum to a playwright, and likens ELIZA to an artificial actor bringing to life his main dramatic character, the Rogerian psychoanalyst [36].

Other examples in the brief history of expressive AI reinforce the notion that the ELIZA effect plays a role in the effectiveness of a piece. Consider *Terminal Time* [13], a work of expressive AI by Michael Mateas, Steffi Domike and Paul Vanouse, in which a computer "constructs ideologically-biased documentary histories in response to audience feedback." A cinematic experience, Terminal Time queries its audience with ideological questions throughout its screening, constructing in real time a documentary film on the history of the world that reflects the audience's responses. Part of the effectiveness of the piece hinges on the audience accepting the fact that the machine is "really constructing" custom documentary films on the fly, as if it were a human editor with insight into the audience's desires and intentions. Though the audience may know the machine is doing nothing more than employing a series of clever algorithms to create poetic cinema, by suspending disbelief and willfully succumbing to the ELIZA effect's siren song, the piece becomes more engaging.

Simon Penny's *Petit Mal* [42] offers another example of how the ELIZA effect plays a role in audience engagement with expressive AI. As Penny writes, Petit Mal is a robot that is "in some sense anti-robot," a robot designed with no particular utility other than to have a particular "charm," to autonomously exist and express itself through its own unique behavior. Aesthetically simple – a slender tube-like body suspended between two bicycle wheels – Petit Mal roams its environment on a kind of curious “voyage of discovery,” reacting and responding to other creatures (i.e. people) as it encounters them. Though Petit Mal exhibits itself as uniquely robotic, with properties and behaviors that are neither “anthropomorphic nor zoomorphic,” one can argue that a degree of anthropomorphism nevertheless influences an audience's perception of the piece. To claim, for instance, that a robot appears to be curious or charming is in some sense to anthropomorphize it. Beyond this, by adopting Janet Murray's [36] broader definition of the ELIZA effect not as unconscious anthropomorphism, but instead as “the human propensity to suspend disbelief in the presence of a persuasive dramatic presence,” there can be no doubt that the ELIZA effect exerts profound influence in an audience's acceptance of Petit Mal.

David Rokeby's *Giver of Names* [48] provides yet another example of expressive AI worth examining in this context. A computer system that, when presented with physical objects on a pedestal, examines them using machine vision and outputs poetic verse-like “names,” Giver of Names is a curious meditation on what machine umwelt could feel like. As it examines objects, the system intentionally gives viewers a glimpse into its own “state of mind” through an accompanying video screen that displays collages of colored words and phrases. The words and phrases represent a kind of poetic metaphorical riff on the object under examination – insight into the machine's own understanding of the meanings of those objects. Though its verbal output is primarily non-sensical, reflecting an experience of the physical objects primarily “alien” to our own, like Petit Mal the system nevertheless exhibits its own peculiar charm and character. Though completely disembodied – being little more than a machine, a screen and a camera – the Giver of Names nevertheless dominates its installation space with its own unique "persuasive dramatic presence."

Finally, consider The Uncanny Dream Machine itself. During The Uncanny Dream Machine's debut at The Beall Center for Art and Technology at the University of California, Irvine, I observed that few, if any, viewers of the piece believed the machine to be actually dreaming. Nevertheless, this was not the desired effect; I had no intention of fooling an audience into believing in a real dreaming machine. Instead, those who appeared to most enjoy The Uncanny Dream Machine were those who engaged with the theatrics of the situation – a cushy chair, an antique reading lamp, and an old-fashioned radio recounting strange and curious bedtime stories.

3.5 Revisiting the Expressive Turing Test

The Expressive Turing Test, a variation of the Turing Test first described in the introduction, offers a potential means of measuring the ability of a work in expressive AI to invoke the ELIZA effect in a viewer. Works of expressive AI that readily and easily engage viewers to suspend their disbelief – that most capably exploit the ELIZA effect for their benefit – score high on the Expressive Turing Test (ETT). Works that fail to persuade a viewer in this way score low. Thus, unlike the real Turing Test, the ETT is not characterized as a binary pass or no-pass situation, but instead measures the effectiveness of a piece on a continuous, and subjective, spectrum. The effectiveness of a piece to invoke an emotional response may fluctuate from interaction to interaction and from person to person. What draws one viewer to suspend his or her disbelief may irritate the next and destroy any potential for the illusion. And whereas the real Turing Test attempts to measure objective factual reality, the ETT attempts to gauge the strength of a subjective emotional relationship.

There are limits to the ELIZA effect's effectiveness. Wardrip-Fruin [58] notes that oftentimes the enchantment of the illusion quickly vanishes once one begins to explore the limits of a particular system. Weizenbaum, for instance, constrained his implementation of ELIZA such that it sustained its illusion of realism only within a narrow domain of conversation and questioning. Any extended interaction with ELIZA soon revealed the weakness of the system, as it could be probed or forced to make inappropriate, non-logical responses by simply overstepping the boundaries of its conversational scope. In the same vein, though The Uncanny Dream Machine can continually generate hundreds of new and unique dream stories, if one were to listen to the babbling of the machine long enough, one would soon discover distinct patterns of repetition. Wardrip-Fruin points out that this kind of breakdown of the ELIZA effect allows viewers to gain knowledge about the otherwise hidden processes of expressive systems. Ask ELIZA the right questions, and over time you may soon puzzle out how the system formulates its responses. Listen to The Uncanny Dream Machine for any extended period of time, and you may soon predict the format and subject-matter of its next dream. In this way, curious viewers can “game the system,” gaining insider knowledge that allows them to either willingly perpetuate or “further compromise” the ELIZA effect illusion [58].

One means of avoiding the negative consequences of an ELIZA effect breakdown involves willingly encouraging it to happen. A system that intentionally exposes aspects of its internal processes while still maintaining expressivity also maintains its potential to serve as a focus of intrigue. There are two challenges with this approach: First, as Wardrip-Fruin explains, "... a system prone to breakdowns is only as interesting as the shape of the processes that the breakdowns partially reveal." [58] Second, appealing to technical matters as a way of sustaining interest in a piece severely limits its audience. Such a tactic relegates the appeal of a work to only a small subset of computer scientists, hackers and sundry technicians.

3.6 Evaluating Computational Narrative

Many works in expressive AI involve the computational generation of narrative. The Uncanny Dream Machine draws from and finds itself situated in a long history of endeavors that explore story as a viable domain for the application of computational tools. Computational narrative is complementary to, but not a subset of, expressive AI. While some works of computational narrative, including The Uncanny Dream Machine, focus on expressivity, others seek to use computation as a means of exploring narrative structure, creativity and linguistics. In many of these latter instances, narrative is presented in a one-sided, non-interactive fashion. Much work in computational narrative results simply in a device that prints out story transcripts. Turner's MINSTREL [55], for instance, is a computer program that generates and prints out short stories in the form of medieval fairytales. Pérez y Pérez's MEXICA [43] prints out stories in the form of Aztec legends. This means that tools for evaluating expressive effectiveness, like the

Expressive Turing Test, must be applied to works of computational narrative with a degree of caution. In interactive expressive works, the ELIZA effect may occurs as a user engages with a piece in real-time. With computationally generated text, however, the ELIZA effect is irrelevant, as a reader will judge the believability, richness, plot, detail and other factors of a story irrespective of its source of authorship. Only when a reader reflects on what a computerized "author" was "thinking" when it conjured up a particular scene or passage does the ELIZA effect come into play. By suspending disbelief we can make believe, for instance, that a computerized author knowingly intended a particular bit of foreshadowing in order to enhance the opening of a story. In reality, however, the foreshadowing occurred only as the result of a clever bit of human programming.

Thus, the ELIZA effect plays only one part in evaluating works of computational narrative, and even then only under specific circumstances. To better gauge the effectiveness of a system of computational narrative, one should first look at the history of the field.

Most histories of computational narrative begin with James Meehan's [33] 1976 thesis work, TALE-SPIN, a software program that generated unique short stories in the form of Aesop's fables. A work of expressive AI only secondarily, TALE-SPIN focused on demonstrating how problem-solving techniques could be applied to story generation. Though it proved early on that computers could be used to generate coherent short stories, many later criticized the system because of its rigid, predefined story structure and its inability to differentiate between character and authorship goals [43].

Turner's MINSTREL fairytale generator follows TALE-SPIN in the historical canon. Unlike TALE-SPIN, which generates stories by solving problems, MINSTREL attempts to model the human creative authorship process through computational means. In doing so, it breaks down the task of authoring a story into twenty-one goals of four main groups. Pérez y Pérez describes these groups as: "thematic goals (whose purpose is to select and develop a theme), drama goals (whose purpose is to create scenes to include in the story suspense, tragedy, foreshadowing and characterization), consistency goals (whose purpose is to keep the consistency of the story in progress) and presentation goals (whose purpose is to present the story in English)." [43] In this way, MINSTREL expanded on TALE-SPIN by accounting for many of its predecessor's prior criticisms. It does so by accounting not just for character- or plot-level goals, but also by modeling the higher-level creative goals a human author may pursue when composing a story. This results in stories that are of richer quality than those written by TALE-SPIN, but still limited to a constrained genre and style.

Many other artificial storytellers soon followed MINSTREL, including Pérez y Pérez's MEXICA, Mueller's DAYDREAMER [35], and Bringsjord and Ferrucci's BRUTUS [6]. All attempt to model human creative processes to achieve their ends. In this, they unfortunately fall far short of real human creativity. To model creativity requires a fundamental understanding of what it means to be creative. Yet, creativity itself often eludes accurate description. We understand creativity in our gut and in our intuition, but only vaguely in our in our intellect. Margaret Boden's *The Creative Mind* [4] offers perhaps the most sophisticated account of the phenomenon, dividing it into two distinct senses: psychological creativity (P-creativity) and historical creativity (H-creativity). She writes:

> P-creativity involves coming up with a surprising, valuable idea that's new *to the person who comes up with it.* It doesn't matter how many people have had that idea before. But if a new idea is H-creative, that means that (so far as we know) no one else has had it before: it has arisen for the first time in human history.

In short, P-creativity involves the rearrangement, juxtaposition or assimilation of existing ideas into something historically novel for the creator, whereas H-creativity requires the generation of something entirely and uniquely new for the world. Personal and global historical novelty are the keys to Boden's definitions, as she rejects the notion that creativity involves mere novel recombination of ideas. Instead, her emphasis concerns novel recombination which not only did not happen before, but also *could* not have happened before. For instance, the act of my writing the previous sentence would not qualify as P-creative because though I may have never written it before, I *could have.*

Drawing on Boden's nomenclature, Pérez y Pérez introduces a third type of creativity: computational or C-creativity. A C-creative computer system is one that "generates knowledge that does not explicitly exist in the original knowledge-base of the system and which is relevant to produced output." [43] C-creativity offers a complementary evaluation strategy to the Expressive Turing Test – one that focuses on the particular facets of computational narrative the ETT fails to address. Specifically, it offers a means for evaluating effectiveness through story predictability, defined by Pérez y Pérez as the degree to which the output of a computerized storyteller can be predicted when the content of the system's knowledge-structures are known. The term “output” implies both content and structure. A highly C-creative storytelling system generates stories in which both the plot and mechanisms used to convey the plot evolve and remain unpredictable between storytelling episodes. The relationship between C-creativity and story predictability is thus inverse; a highly C-creative system exhibits low story predictability, and vice-versa.

Pérez y Pérez suggests a possible line of questioning to aid in the assessment of story predictability for a story generation system. In particular, he suggests analyzing:

- The type of knowledge stored and method of knowledge representation employed within the system.
- The mechanisms used to update the system's knowledge.
- The mechanisms used to ensure interestingness, novelty, story coherence, and proper natural language structure.
- The mechanisms used to form a story from the information in the knowledge-base.

Here, Pérez y Pérez exploits the fact that vast knowledge databases lie at the heart of nearly every story generation system so far conceived. Story generators require human programmers to inform them of what we take for granted as simple common sense and common knowledge. They also require human programmers to equip them with knowledge of characters, character capabilities, the facets of possible story worlds, and possible plot and authorship goals. Furthermore, story generators by themselves have no knowledge of natural human language or vocabulary; this too must a human programmer outfit the system with in order for it to generate anything of interest to a human audience.

Thus, when we begin to tease apart the inner workings of a storytelling system, when we probe into its inner knowledge structures the way Pérez y Pérez suggests, or when we intentionally break the ELIZA effect illusion, much of the "magic" of the system disappears. By evaluating systems of computational narrative in these ways the stark limitations of the entire enterprise become apparent. What we find are not creative systems, but creative human engineers of systems that in turn *appear* to be creative – clever prestidigitation from ELIZA to TALE-SPIN to BRUTUS.

3.7 Challenges to To Machine H-Creativity

This type of investigation ultimately begs the question: Could a computer ever *really* be creative? In *The Creative Mind*, Boden concerns herself primarily with the phenomenon of human creativity; she is interested only tangentially in the possibility of true machine creativity. Nevertheless, she does discuss it, citing a famous quote by Ada Lovelace, friend and colleague of Charles Babbage, inventor of the mechanical precursor to the modern computer, the Analytical Engine. In a retrospective on her work with Babbage, Lovelace famously declared: "The Analytical Engine has no pretensions to *originate* anything. It can do *whatever we know how to order it* to perform." [4][54] Framed in Boden's terms, one could say that Lovelace implied that computers could never be H-creative, but that they possibly could, in a sense, be P-creative.

Boden addresses Lovelace's assertion about computational creativity in four parts:

1. Can computational ideas help us understand how *human* creativity is possible?
2. Can computers, now or in the future, ever do things which at least *appear* to be creative?
3. Can a computer ever *recognize* creativity?
4. Can a computer ever *really* be creative (as opposed to an apparently creative disembodied zombie)?

Boden concerns herself primarily with only the first three questions, answering yes to them all, but it is the fourth that is of greatest concern here: Can a computer ever *really* be creative? Boden gives a tentative and muddied no. Bringsjord and Ferrucci, creators of the BRUTUS story generation system, give a more resolute no. Instead, they agree with Lovelace, concluding that: Yes, computers can be P-creative. But no, they cannot ever be H-creative. [6]

The challenge with Boden's definition of creativity is that it makes it too easy to ignore the fact that creative action is an embodied, situated and planned act of meaning making. It fails to emphasize that the output of some creative action must have meaning for the creator if not also for an intended audience. To randomly dribble with colored paints on a canvas and call it art involves absolutely no creativity; to dribble with colored paints on a canvas in a novel, meaningful and intentional way, on the other hand, is arguably creative. For this reason, I propose a clarifying amendment to Boden's definition, arguing that to qualify as creative, the output of some action must not only be historically novel, but must also convey or contain meaning for the creator.

Consider the following thought experiment: Suppose I create a computer program that re-combines portions of an existing collection of hundreds of culturally-significant short stories – fables, myths and legends – into novel short stories that are equally engaging, entertaining, and have nearly undetectable traces of their original components. Suppose I create another system, a computer-driven paint plotter that algorithmically spurts various colors of paint onto a canvas to create novel paintings in the style of Jackson Pollock. Is either of these hypothetical computer programs being creative? Some, like Bringsjord and Ferrucci, may argue that the former system is creative – at least P-creative in its systemic historical novelty – while the latter system constitutes nothing but mimicry or sheer random nonsense. Perhaps the differentiation here asserts itself because stories automatically contain some intrinsic meaning for us as humans, whereas random paint splatters typically do not. Boden may argue that both systems could be P-creative *if* they self-modify in such a way that after each session they produce output that they could not have produced in the past. I argue, however, that neither system is creative because neither system understands either its actions, its output or its supposed goal. To each computer, all portions of the respective activity – story writing or painting – constitute nothing but sheer random nonsense because neither has a sense of the activity's meaning. As Lovelace insinuated and Searle affirmed, computers manipulate symbols without the slightest idea of what they're doing. This is what is known in cognitive science as the "symbol grounding problem." [21]

3.8 Full Circle on Embodiment

Searle offered one example of the symbol grounding problem when he penned the Chinese Room Argument. Just as Searle, sealed off in a room, can appear to understand Chinese symbols with nothing more than a sophisticated rulebook, so too can a computer system appear to understand its myriad inputs and outputs when in reality it hasn't a clue. Neither "system" accounts for semantics; instead, we as observers insert the semantics unwittingly. The problem arises for precisely the same reason that we find creativity in the hypothetical story generator, but not in the paint splattering apparatus: we are too easily beguiled by the meaning we automatically find in familiar linguistic symbols. Wittgenstein [60] made this point abundantly clear when he described how the possession of a word for a concept often deludes us into thinking we have the concept itself. As such, it is worth reiterating that any meaning we find in linguistic symbols is not intrinsic to the symbols or correlative symbol systems themselves. Instead, as Steven Harnad [21] writes, any meaning we do find will "be parasitic on the fact that the symbols have meaning for us," deriving entirely from the "meanings in our heads."

Of the many schemes proposed for solving the symbol grounding problem, the most promising poses a significant challenge for the entire enterprise of constructing truly creative computational systems. This is the suggestion that conceptual knowledge – the "meaning" represented by the symbols – is embodied [17]. Under an embodied approach to symbol grounding, meaning derives not from anything intrinsic to symbolic systems themselves (which the Chinese Room Argument shows cannot be so), but instead from lived experience with a sensory-motor system. When Harnad claims that the symbolic output of a computer only has meaning for us because it is "parasitic" on the "meaning in our heads," he does not immediately ask exactly where the "meaning in our heads" comes from. Ironically, at least according to the embodied approach to symbol grounding, the "meaning in our heads" actually comes from our bodies.

Basing their argument on neuroscience research which shows how acts of *imagining* and *doing* share the same neural substrate, Gallese and Lakoff [17] propose that "the sensory-motor system … characterizes the semantic content of concepts in terms of the way that we function with our bodies in the world." For instance, they suggest that the meaning behind the word grasp derives from "our ability to imagine, perform, and perceive *grasping*." When we imagine grasping, we invoke the exact same motor circuits we would if we were actually performing it. Thus, in order to even imagine or understand a concept, we first require the sensory-motor circuits that allow us to do it. The ultimate conclusion: semantics originates in the body.

We return again to the challenge of computer embodiment. Just as a computer cannot dream because it lacks a body, so too can no computer ever become the next Shakespeare for the same reason. Not only do all computational narrative systems lack the ability to understand their own stories – and thus lack true creativity – as disembodied zombies they also lack conscious subjectivity. Few authors would argue against the idea that the weaving of a compelling tale requires assuming the point-of-view of one's characters. Bringsjord and Ferrucci, for instance, quote Henrik Ibsen [6]: "Before I write down one word, I have to have the character in mind through and through, I must penetrate into the last wrinkle of his soul … Then I do not let him go until his fate is fulfilled." Yet, not only can a computer not understand its characters, it cannot assume any of its character's points-of-view. Furthermore, it cannot assume any point-of-view, even its own. There will never be a computerized Ibsen as long as there is never a something-it-is-like-to-be a computerized Ibsen.

3.9 Criticisms of the Argument Against Machine Creativity

Boden claims that the embodiment argument against computational creativity suffers from a bias towards "poetic" forms of creativity. She writes, "A computer's alien indifference to winter weather may jeopardize its poetic sensitivity, but not its originality in science and mathematics." [4] Yet, by reframing creativity as more than originality, and instead as an act of meaning making, the argument against computational creativity still stands. A computer can be original in science and mathematics – it could uncover new theories in particle physics, for instance – but without a human to decipher the meaning and value of these theories, its originality is simply number crunching.

Others may argue that though embodied experience may be necessary for proper creativity, this does not preclude synthetic bodies. Some may argue that a robot, for instance, could be creative. Using enactivism as a theoretical foundation, teams of roboticists have successfully constructed "embodied" machines capable of navigating around, interacting with, and learning from their environment [7]. Through his zombie gedanken-experiment, however, Searle offers one means of demonstrating how even a robot, though embodied, still does not possess the subjective conscious awareness necessary to understand the meaning of its actions. It could be a zombie.

Another challenge with robotic embodiment is that it is not true embodiment in the enactivist sense of the word. To *have* a body is not to *be* embodied. Because of the static nature present technology, a robot's structural coupling with its environment remains artificial. It does not evolve and change with its environment such that its embodied experience grounds its understanding of the world. Instead it is programmed, engineered, built and inserted into a pre-existing world. Any artificial bodily coupling relies on the robot's human creators to recognize the necessity for modification and then act accordingly. Such coupling is not self-regulating nor self-initiated.

Enactivism borrows its definitions of embodiment and structural coupling from Maturana's [32] earlier theory of autopoeisis. Autopoeisis describes living, embodied, biological systems as continuously regenerative and self-creating. No technology to date offers the capability to build similar regenerative and self-creating systems, which, according to enactivism, is what bodies are. Thus, under enactivism, presently only living biological systems can be embodied systems, and only living biological systems can have subjective conscious experience or the capacity for real creativity.

Nevertheless, technology constraints evolve and broaden. What is inconceivable today may not be so tomorrow. A synthetic technology may one day arise that allows humans to build living, embodied, autopoietic systems out of non-biological components – systems capable of real creativity and real, original storytelling. But, would we call such systems "machines?" Would we call them "computers" or "robots?" In all likelihood, we would not.

3.10 Machine Creativity as a Worthwhile Pursuit

At present, storytelling machines remain disembodied zombie storytelling machines. They remain a medium for expressing *human* creativity and commenting on *human* issues through programmed autonomous behavior. Through this lens the quest to build creative machines transforms from a hopeless folly to a worthwhile pursuit. The caveat that such machines will not ever achieve true creativity becomes moot.

This is the alternative that expressive AI offers: a means of applying the rich set of technical practices found in traditional AI towards works of cultural production. As works of expressive AI, storytelling systems and dreaming machines are not scientific tools for studying neurobiology or computation or even creativity. Instead, they are cultural artifacts that ask us to reflect on our own experience of what it means to be alive and human.

Chapter 4: Technical Documentation

4.1 Introduction to the Technical Documentation

The following chapter provides an in-depth technical description of The Uncanny Dream Machine, its design, architecture, and inner workings. Here, I present the technical details of The Uncanny Dream Machine without extensive theoretical commentary, reserving such discussion for the final chapter. I do, however, occasionally delve into particular technical design rationale, as well as describe hurdles or obstacles overcome in the production and exhibition of the final piece.

4.2 Architecture

From a technical perspective, The Uncanny Dream Machine consists of two major components: the dream server – a web services-based AI system written in the Ruby programming language – and the dream machine itself, a highly modified 1940 Philco wooden radio. The client-server architecture employed in The Uncanny Dream Machine serves two primary purposes: First, it off-loads the heavy computation required by dream generation to a computer sitting "behind the scenes," allowing the machine itself to retain a small physical footprint. Second, it affords significant flexibility in the way future iterations of The Uncanny Dream Machine may be presented. For instance, though the machine as presented in its first exhibition offered only a single-user storytelling experience, a server-based architecture permits future iterations to offer multi-user or virtual experiences that target a wider audience.

Server-machine communication occurs over 802.11g wireless ethernet, commonly known as "Wi-Fi." This ensures that the machine can be placed an arbitrary distance from the server without unsightly cords interfering with aesthetics or experience. The system's Wi-Fi communication employs a closed network, such that the server and the machine are the only devices on the network, communicating in a secure fashion. For the gallery installation, this decision was made primarily to prevent possible malicious or mischievous tampering with either component of the system.

All of the dream generation work of the system occurs on the server. This includes maintaining the database of recorded dreams, analyzing and “re-mixing” the recorded dreams into novel dream narratives, and converting the dream narratives into synthesized audio files for streaming to the machine. The machine's responsibilities are relatively light in comparison; these include reading channel selections via a control knob, selecting the appropriate server stream based on the current channel selection, and visualizing the volume level of the audio playback by controlling the brightness of a multi-colored LED.

4.3 The Server

The server is programmed using the Ruby programming language and the Ruby on Rails [20] web framework. I chose Ruby and the Ruby on Rails framework for two reasons: First, I was already intimately familiar with both, allowing me to build, experiment with and test the system quickly without the additional challenges of learning a new programming language or framework. Second, the Ruby language happened to be well-suited for linguistic parsing and analysis, second perhaps only to Python and its extensive Natural Language Toolkit (NLTK) [3].

A substantial and crucial component of the server is its database, implemented as a standard structured query language (SQL) relational database using MySQL. The server uses its database to maintain a collection of recorded human dream transcripts (in this case, six months of my own dream transcripts), as well as a collection of linguistic and semantic statistics about those transcripts. After recording, the system analyzes each transcript to generate such statistics, which in turn provide the necessary fodder for dream narrative generation.

Dream generation also relies on two other static third-party databases, not generated nor maintained by the system. The first of these databases is the Princeton University WordNet database, a vast lexical database of the English language compiled under the direction of George Miller [34]. WordNet offers an extensive ontology of English words, organized in a hierarchical agglomeration of hypernyms (words whose semantic fields include other words), and hyponyms (words whose semantic fields are included by those of other words). For instance, a WordNet hypernym search on the term "redwood" results a series of terms that grow semantically broader at each hierarchical level: cypress, wood, plant material, material, etc. Figure 4.1 shows a portion of the complete hypernym hierarchy for one sense of the word redwood in the WordNet database. In addition to hypernym/hyponym hierarchies, WordNet also permits the lookup of word synonyms, antonyms, holonyms, and meronyms, depending on the queried word's respective part of speech.

```
=> redwood
     => wood
          => plant material, plant substance
               => material, stuff
                    => substance
                         => matter
                              => physical entity
                                   => entity
```

Figure 4.1 Hypernym Hierarchy for the Word "Redwood"

The second third-party database employed The Uncanny Dream Machine server is the University of South Florida Database of Free Association Norms (DFAN) [40]. DFAN is a collection of over three quarters of a million responses to nearly twenty years of psychoanalytic free association testing. During free association testing, analysts present human test subjects with cue words, typically displayed on a series of notecards. As each word is presented, test subjects are asked to write down the first word that comes to mind in response to the particular cue word. For instance, the cue word "cheddar" may lead a test subject to write down the target word "cheese." DFAN includes statistical data on how frequently a particular cue word leads a test subject to write down a particular target word, known as cue-to-target strength. The cue-to-target strength of "cheddar" and "cheese," for instance, is 92 percent, meaning that 92 percent of test subjects are likely to write down the word "cheese" when presented with the word "cheddar."

Unfortunately, neither of these databases offer easy or immediate data access using the Ruby language. Natively accessing WordNet via Ruby required the creation of a custom library of code for parsing WordNet's indented textual output into navigable tree data structures. And natively accessing DFAN required translating its lengthy statistics tables into a series of SQL statements that could then inserted into the MySQL database for querying.

The server exposes these databases and other functionality as a series of representational state transfer (REST)-based web services, transferring data either in XML format, or in the case of audio streaming, as MPEG layer-3 (MP3) audio. Novel dream narratives are generated by the server as requested by users in real-time, saved as temporary text files, converted to WAV audio using the Cepstral [9] text-to-speech software, and finally encoded as MP3 audio using the open-source LAME MP3 audio encoding software.

4.3 The Machine

The physical machine itself may be likened to a kind of customized Wi-Fi radio. The wooden body of a 1940 Philco radio houses an ASUS WL-520GU wireless router running the OpenWrt Linux [41] operating system. The router uses an external SYBA USB sound card for audio playback through a standard computer speaker, the sound source being an MP3 wirelessly streamed from the server using the Linux music player daemon (mpd).

Bakelite knobs on the front of the machine connect to potentiometers, the first of which controls the volume of the embedded computer speaker, and the second of which connects to an internal microcontroller for input processing. The microcontroller is an 8Mhz Atmega328P, embedded on an Arduino Pro Mini [2]. This microcontroller serves two primary functions: First, it reads the dream channel selection from the position of the potentiometer, notifying the wireless router of the new channel selection, and updating the color output of a multicolor red-green-blue (RGB) LED visible behind frosted glass on the front of the radio. Second, it reads the volume level of the current audio playback and adjusts the brightness of the RGB LED accordingly, such that the LED fades and brightens with the volume of the synthetic storyteller's voice. Figure 4.2 shows a complete schematic diagram of the volume monitoring and LED control unit for the machine.

The machine visually delineates each possible dream channel selection as a color on the RGB LED. I chose a color for each dream emotion that most closely matched its corresponding color on Plutchik's Wheel of Emotions [45] (see Figure 4.3), while simultaneously maintaining the highest level of brightness and contrast possible within the LED's luminance range. In the final implementation, curiosity matched with yellow, frustration with orange, shame with cyan, fear with green, anger with red, joy with white, worry with blue, and embarrassment with magenta.

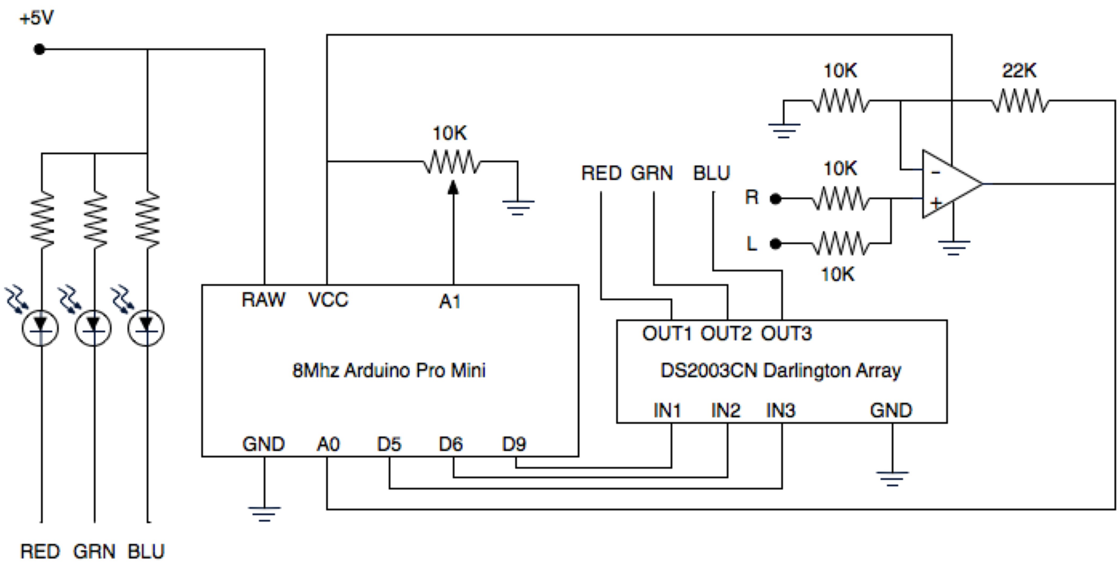

Figure 4.2 Schematic of LED Control Unit

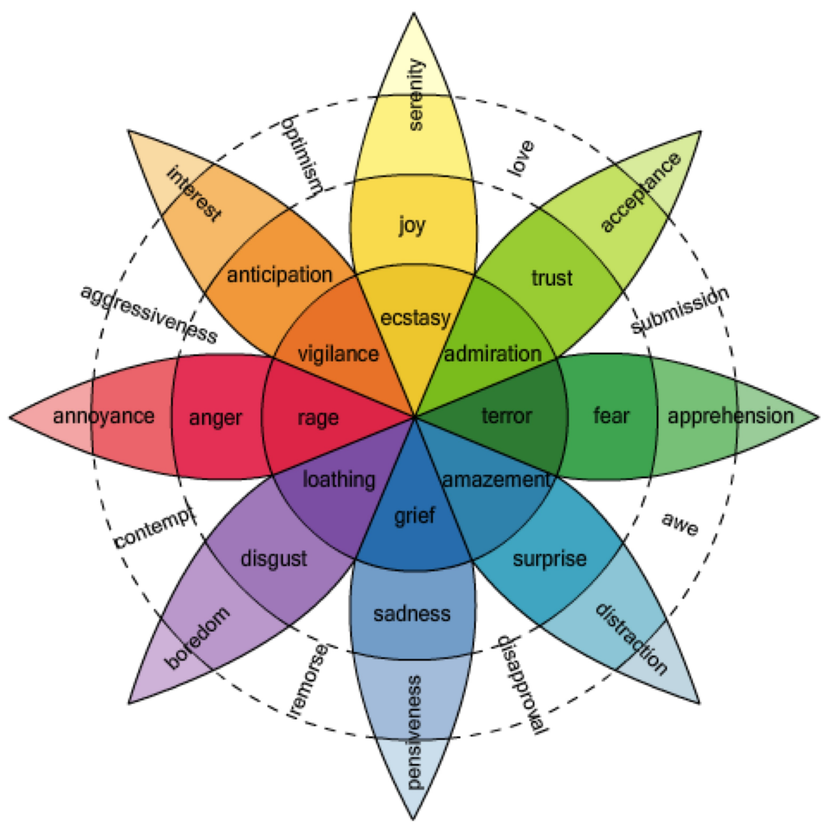

Figure 4.3 Plutchik's Wheel of Emotions

4.3 Dream Transcript Analysis

To generate dreams of an uncanny nature requires that the system have significant semantic and stylistic information about its corpus of real human dreams. In order to acquire this information, The Uncanny Dream Machine server performs semantic and statistical analysis on each dream transcript as it is entered into the system. Dream transcript analysis proceeds as follows:

1. The server first uses a Brill part-of-speech tagger [5] to tokenize and label each word in a dream transcript according to its part of speech – noun, verb, pronoun, adjective, adverb, etc.
2. Significant words – namely, nouns and verbs – are then lemmatized and recorded in the database along with their respective usage frequency. Adjectives are also recorded with a pointer to the noun that they modify in context. For instance, if the parser encounters the phrase "old socks," the adjective "old" is recorded as a modifier of the noun "socks." Words that are not found in the WordNet database, excluding proper nouns, are thrown out and not recorded.
3. Next, the system uses the WordNet database to determine, for each noun, whether it describes an actor, location, or object. For instance, when looking up the noun "artist," the system will find that a hypernym for "artist" is "person," thus categorizing "artist" as a kind of actor. Unfortunately, this particular methodology is also prone to failure. For instance, a "horse" is both a mammal and a kind of gymnastics equipment, meaning that it could be categorized as

either an actor or an object depending on context. The present version of the system employs a rather dumb strategy in such cases, choosing the first sense of the word regardless of contextual placement. Future iterations may employ a more sophisticated word sense disambiguation algorithm to improve on this process.

4. Finally, the system uses the WordNet database to perform semantic relationship graph (SRG) analysis [18] on the transcript using all significant nouns not thrown out in prior steps. SRG analysis offers a means of determining the semantic, or conceptual, content of a particular dream transcript without human intervention. For instance, a dream transcript that includes the words, "crash," "explosion," "siren," and "police" may lead the system to attach the terms "mayhem" or "tragedy" to the transcript, even though neither may appear in the actual textual body. Figure 4.4 shows a portion of a sample semantic graph of one of my own dream transcripts.

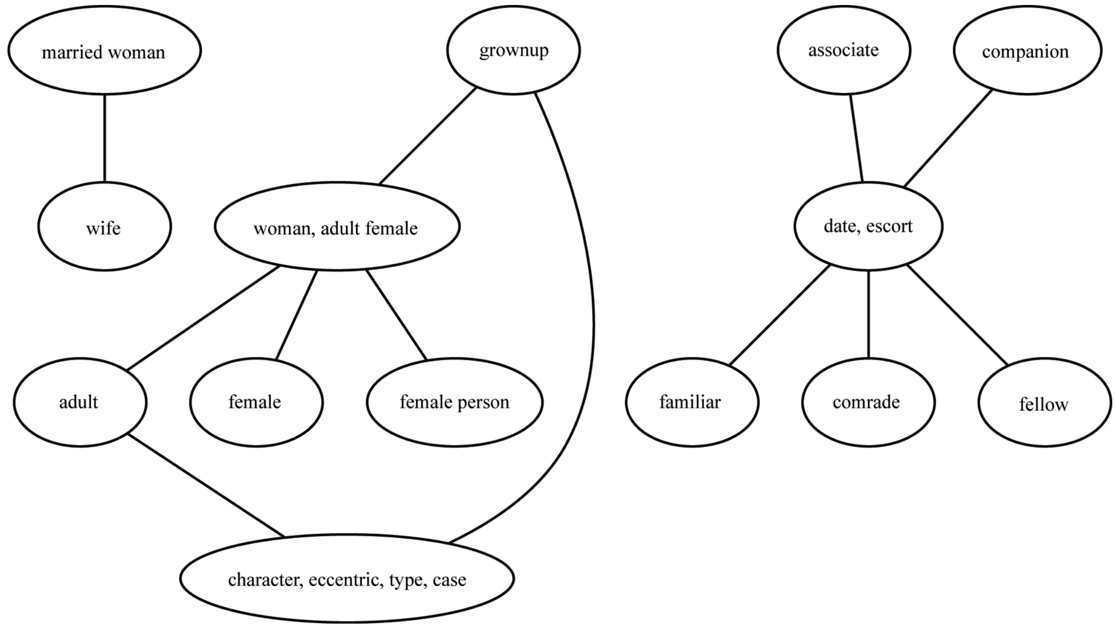

Figure 4.4 Semantic Relationship Graph

Though surprisingly accurate in a number of cases, SRG analysis also frequently fails to produce accurate results with dream transcripts that include a number of disparate, but adjacent, concepts. SRG analysis works best with conceptually homogenous texts, such as newspaper articles, but often flounders when presented with the bizarre and heterogenous. For this reason, the system extracts only the most significant conceptual terms found in each graph, determined by the number of edges incident to each term's corresponding vertex in the graph. High-degree vertices are deemed semantically more important than those of low-degree. In the sample graph displayed in Figure 4.2 of a dream about attending a wedding, the system would extract the terms “woman” and “date” as most semantically relevant.

The system maintains records of all semantic concepts associated with each dream transcript, as well as concept usage frequency across transcripts. To aid even further in the subsequent dream generation process, a human, preferably the one who experienced the dream, must categorize each dream transcript with a word describing its most dominant emotional quality – fear, curiosity, anger, shame, frustration, etc. Though it may be possible to devise a WordNet-based algorithm similar to SRG analysis that permits automated emotional categorization based on verb and adverb usage, I determined that such research was beyond the feasible scope of the current implementation. Future iterations of The Uncanny Dream Machine may improve on this method.

4.4 Dream Narrative Generation

The dream narrative generation process begins with the viewer's selection of a desired emotional quality. Once an emotion has been chosen, the system looks up all concepts related to the emotion that were extrapolated during semantic relationship graph analysis. Two concepts are chosen at random based on their usage frequency in the corpus. For instance, if the concept of “mayhem” appears in fifty percent of all recorded nightmares, but the concept “motherhood” appears in only twenty-five percent of the recorded nightmares, then “mayhem” is twice as likely to be chosen as “motherhood” in all fear-based dreams. Along with the emotional quality, these concepts guide the subsequent dream creation process, narrowing the selection of people, locations and objects that will appear in the final narrative to those most appropriate and emotionally relevant.

To generate the final style and sentence structure of the narrative, The Uncanny Dream Machine expands on the methods described in Bulhak's “On the Simulation of

Postmodernism and Mental Debility using Recursive Transition Networks." [8] Namely, it uses recursive transition networks (RTNs), composed in the form of context-free grammars, to produce grammatically correct and semantically reasonable texts. An RTN is a directed acyclic graph, the nodes of which describe tasks to perform, and the edges of which point to the next task in a series of tasks. For sentence construction, RTN nodes describe the possible fragments of text that can be inserted into a sentence at a particular location. For instance, a sentence may begin with a preposition, followed by an optional adjective, followed by a noun, and then a verb. Figure 4.5 shows an RTN for this type of sentence generator written in the form of a grammar.

```
sentence ::= preposition [ε | adjective ] noun verb
preposition ::= "the" | "a"
adjective ::= "green" | "blue" | "furry"
noun ::= "dog" | "cat
verb ::= "runs" | "sleeps"
```

Figure 4.5 A Simple Recursive Transition Network

The grammarized RTN described in Figure 4.5 generates rudimentary sentences of the form, "the green dog runs" or "a furry cat sleeps," grammatically valid but not particularly praiseworthy pieces of prose. Nevertheless, The Uncanny Dream Machine employs this same method in a more sophisticated manner, generating RTNs as needed in real-time from the statistical and linguistic data stored in its MySQL database. For instance, if a grammar calls for an actor noun – that is, a noun which describes an organism or entity that can perform an action – The Uncanny Dream Machine will query its database for all appropriate nouns and choose one at random based on its usage frequency and relationship to the desired emotion. If the system is tasked with generating a joyful dream for example, and it requires an actor, it will query all actors associated with joy, create a frequency distribution of the results, and insert a random selection at the appropriate location. Thus, the more frequently a person or other actor appears in the real dream transcripts in the dream corpus, the more likely that same person or actor will appear in generated dreams.

In addition to using words extracted from the transcript corpus and stored in the local database, the system also makes extensive use of WordNet to look up relevant synonyms, hyponyms and hypernyms. This technique adds significant linguistic variety to the final output, allowing it to employ concepts and vocabulary related to, but not found in the original corpus. For instance, if the system originally chooses the word "artist" to fulfill the actor role in a new dream, it may instead substitute the hyponym "Marcel Duchamp" in the final output. In this way, generated dreams remain conceptually similar to, but also uniquely different, from those on which they are based.

This strategy requires painstakingly hand-programming a large collection of generalized recursive transition network templates. Though similar in nature to the RTN of Figure 4.5, these templates differ in their broad generality, employing special character patterns that the system can use to insert actor words, location words, objects, adjectives, verbs, hypernyms, hyponyms, and so forth. The templates themselves read like vague story frameworks, with sentence generation patterns that look much like the following: `#Actor1# #Actor1Action# Preposition #ArticleizedObject#`

`Conjunction #Actor1Action# #Actor2#`. For each variable in a sentence generation pattern, the system substitutes an appropriate word, chosen randomly by its usage frequency in the context of the selected emotion and set of semantic concepts. Separate linguistics algorithms ensure the proper pluralization of and article usage for each word. This process continues recursively until the system has generated a narrative of on average between five and ten sentences, the length of a typical real dream transcript in the corpus.

DFAN is used during dream generation for two purposes: First, it serves as a rudimentary “common sense” knowledge base used to look up appropriate verbs for given actors or objects. For instance, if the system requires a verb to be used in conjunction with the word “dog,” it may query DFAN to get a list of verbs – bark, sleep, walk – related to dog. Though not designed for this purpose, as a large collection of human knowledge encoded in association data, DFAN proves to be a computationally light-weight source of common sense information. This is especially true when compared to other knowledge bases designed for this purpose, such as the Cyc ontology [11]. The downside with this approach lies in the fact that it does not offer a means for determining how related verbs apply in context. Different verbs relate to different nouns depending on whether the noun is the subject or object of a sentence. For instance, though dogs do bark, it makes no sense to “bark a dog.”

DFAN's second role is to serve as a tool for “conceptual jumping.” Mid-narrative conceptual jumping renders the output of the system more realistic and dream-like. In real dreams, events occur as we think them, such that in one moment we may be dreaming about one situation, and in the next moment the situation transforms into something entirely different (though strangely related). Conceptual jumping from cue to target words in DFAN affords The Uncanny Dream Machine a way of simulating this phenomenon in its own narratives. For instance, the system may start with the concept word “mother,” and through cue-to-target jumping in DFAN, end with the concept word “tupperware.” Guiding concepts do not determine the narrative structure of generated dreams, but they do guide noun (actor, object, and location) choice. As such, the conceptual jumping strategy provides another source of subject-matter and linguistic variety.

4.5 Installation

During its opening exhibition at the Beall Center for Art and Technology at the University of California, Irvine, I designed The Uncanny Dream Machine installation such that it would be reminiscent of a quiet reading nook in a living room. The machine itself sat on a small, oak side-table adjacent a cushy chair and reading lamp. The primary goal of this design was twofold: First, it provided gallery viewers with a quiet and comfortable location in which to sit and listen to the machine's many stories. Second, this style of installation dovetailed with the aesthetics of the machine's 1940's-era radio housing, itself chosen to evoke the nostalgia of being lulled to sleep by radio programs during a quiet evening at home.

Chapter 5: Reflections on The Uncanny Dream Machine

5.1 Design Intentions

The Uncanny Dream Machine stands apart from many other prior works in computational narrative in that it moves beyond authorship into live storytelling. While one could extract the software that lies of the heart of the machine into a standalone story generator, one would eliminate the critical oral storytelling component in the process. As such, like Mateas' *Terminal Time*, The Uncanny Dream Machine represents a hybrid approach to computational narrative that draws on standard storytelling techniques while simultaneously including an element of performance.

The narrative aspect of The Uncanny Dream Machine includes a system that analyzes and draws from an existing corpus of narratives. This differentiates the piece from prior works in computational narrative – TALE-SPIN, MINSTREL, BRUTUS – in that it does not attempt to model the creative writing process, nor compose narratives from scratch. Instead, it "re-mixes" existing narratives into new narratives. The purpose behind this design spawned from my initial decision to explore dreams as a narrative genre; I sought uncanniness. In invoking the term uncanny I refer to its Freudian meaning, denoting an experience simultaneously foreign and familiar. An uncanny experience is an experience characterized by a kind of unsettling psychological dissonance – a conflict between intuitive feelings of familiarity, accompanied by an intellectual acknowledgment of the strange. Freud [15] notes that these unsettling experiences often induce fear in the experiencer. With the The Uncanny Dream Machine, however, I sought not fear, but to self-induce a kind of minor paramnesia or déjà vu-like sensation. I sought to create dreams familiar in their characters, locations, and concepts, but strange in the way these characters, locations and concepts are assembled and employed.

The Uncanny Dream Machine also plays upon another meaning of the term uncanny provided by Jentsch [24], an early influencer of Freud's. Writes Jentsch, "In storytelling, one of the most reliable artistic devices for producing uncanny effects easily is to leave the reader in uncertainty as to whether he has a human person or rather an automaton before him in the case of a particular character." Albeit the lack of uncertainty as to the machine's existence as automaton, it nevertheless evokes this notion of uncanniness. The Uncanny Dream Machine is an automaton that emulates uniquely human behavior. Were it to emulate this behavior in a way so believable as elicit ELIZA effect-like responses, it might establish in a viewer uncertainty of its unconsciousness despite overwhelming additional evidence to the contrary. The ELIZA effect thus in some ways relies on the human ability to recognize the uncanny, and The Uncanny Dream Machine plays on this fact. Its dreams are both familiar and foreign, their origins both human and machine.

This approach to the uncanny gives The Uncanny Dream Machine the opportunity to serve as a rudimentary device for the production of false memories. My experience with my own dreams is that they have occasionally induced déjà vu-like or paramnesic experiences in waking life. I have, on occasion, dreamt experiences that days, weeks, or even years later I experienced similarly in my day-to-day reality. This phenomenon has served as a continual source of personal intrigue, and drew me to the idea of constructing a machine that could intentionally induce it. By pre-loading The Uncanny Dream Machine with a collection of my own dreams, I hoped that it could then produce new dream narratives that would be uncannily similar to, but different from, my own. By listening to the generated dreams for an extended period prior to sleep, I hoped to induce confusion as to my own dreams' origins.

This quest for the paramnesic guided much of the ensuing design. First, it liberated me from the dominant paradigm that drives much prior work in computational narrative, namely the idea that to build a storytelling machine one must first model the story writing process. I focused instead on computational methods that would improve the expressivity and déjà vu-like effect of the machine, regardless of their domain of applicability outside the work. For instance, the use of semantic graphs for computational dream analysis is unique, and may spawn further research into how computational analysis of dreams can help us to better understand their significance. The use of The University of South Florida Database of Free Association Norms, both for conceptual jumping and as a common sense knowledge-base, also likely represents a first in the field.

This quest also liberated me from the constraints of traditional narrative structure. By focusing on dreams as narratives, which my their very nature are characterized by the bizarre and non-sensical, I could focus less on the plot and character congruency found in traditional narrative, and more on the linguistics of the final output. This I deemed crucial: that the language match the style and character found in the original dream corpus, and that it invoke the rich and strange imagery common in most dream experiences.

In dreams, plot is secondary. Of primary concern are the people, locations, objects, situations and images that appear within. Most dream narratives, such as those recorded by Freud or Piaget, are narratives in the most rudimentary sense, and are perhaps better described as short situational descriptions. Unlike true stories, dreams do not have beginnings, middles, or endings. Instead, the dreamer finds him or herself already immersed in a situation, accepting it as matter-of-fact, with little to know wondering of how the situation came to be, what happened before, or how it will be resolved. As States [52] writes, "... the peculiar *presentness* of dream temporality is unlike anything in the normal waking state." A dreamer is simultaneously experiencer and omniscient author; simply by thinking something will or will not happen makes it so.

This implies that dream narratives do not have authorship goals of the kind systems like MINSTREL seek to address. The dream unfolds as the dreamer thinks it, with no plan or goal in sight. The narrative generation strategy employed in The Uncanny Dream Machine reflects this fact, first choosing the actors, objects, and physical location of the dream, and then tossing them together into an emotionally relevant situation. It makes not attempt to resolve the situation – as dream situations are rarely, if ever resolved – but simply to describe it.

Finally, the desire to create paramnesic experiences also informed the machine's physical design. I sought to design the piece such that it would look appropriate in a domestic setting, either beside a bed or a comfortable chair. This lead to the appropriation of an antique radio to serve as the machine's housing, and a colored LED diffused behind frosted glass to provide a visual accent with a television-like glow. I sought to create a situation wherein the listener of the machine could comfortably fall asleep to the machine's narrations, allowing the artificial dreams the opportunity to "seep in" and influence the listener's real dreams. This unfortunately proved challenging for one primary reason: even the best synthetic voice does not sound particularly soothing. Though the technology behind text-to-speech synthesis has improved significantly since the allophone-based technologies of the 1980s and 1990s, synthetic voices remain largely monotone and jarring. Nevertheless, after several weeks of working on the machine well into the evening, I soon discovered its monotone narrations inserting themselves into my own dreams, often startling me from slumber. This was not the déjà vu-like effect I originally sought, but proved to be a curious side-effect of extended interaction with the machine nonetheless.

5.1 Subjective Evaluations

As a hybrid work of computational narrative and performance, a comprehensive evaluation of The Uncanny Dream Machine cannot focus on any one of these facets alone. Instead, for expressive AI works of this nature I propose a hybrid evaluation strategy involving both the Expressive Turing Test and Pérez y Pérez's method for evaluating C-creativity (story predictability).

Evaluation using the Expressive Turing Test is a purely subjective exercise. As mentioned previously, what engages one viewer may bore or irritate another. At the opening exhibition of The Uncanny Dream Machine, many viewers engaged with the theatrics of the piece, sitting in the chair and listening intently to the machine's continuous babbling. Many of the narrations the machine produced during this time proved humorous, either for their content or their absurdity. Rather than work to the detriment of the experience, this often kept listeners engaged, and left them wondering what absurd or non-sensical scenario the machine would describe next. For other listeners, however, the monotone synthetic voice proved to be either too irritating or too incomprehensible. The semi-interactive nature of the piece also often proved to be non-obvious. Some touched the knobs on the radio only with extreme trepidation, perhaps because of their apparent antiquity or because the listeners were unsure of their utility.

In addition, bugs and other imperfections often lead to a either a failure to induce, or a complete breakdown of, the ELIZA effect. The challenge of differentiating between subject and object when using verbs suggested by DFAN, for instance, sometimes lead the machine to produce sentences like: “I bark the dog” or “The cake bakes me.” Such instances proved rare, however. A far greater challenge proved to be the production of sentences with semantic content in congruence with common sense. For instance, the machine would often produce sentences like, “I find a blue car in my pants and throw it at my brother.” Fortunately, the inherent bizarreness of the dream world provides some leniency for this type of imagery. It is conceivable that one could dream of having a full-sized sedan in one's trousers. Nevertheless, most dreams, having their basis in the waking umwelt of the dreamer, do not deviate this far from real-world physical possibility.

A final imperfection of the system involved hidden and unresolved syntactic errors in the recursive transition network templates. These errors lead the system to, on very rare occasion, begin spouting out variable names, filenames, and other syntactic gibberish, much to the surprise of the audience. Like the absurd nature of the narratives themselves, this unintended quirky behavior sometimes intrigued listeners even further, drawing them to speculate on the hidden inner workings of the machine. Furthermore, it served as a reminder to the audience that the machine really was constructing its narratives in real-time, and not simply reiterating the same small set of pre-rendered narratives ad infinitum. This is a challenge that nearly all practitioners of expressive AI encounter: the quest to find the delicate balance between how much to conceal or expose of a system's inner workings. Show too much, and the “magic” of the piece vanishes; show too little and the audience may not learn to appreciate the complexity and subtlety of the system's behavior.

From the standpoint of story predictability, Pérez y Pérez's C-creativity criteria are only marginally applicable to The Uncanny Dream Machine. The goal of producing uncanny dreams is in some ways oppositional to the desire to produce unpredictable stories. An examination of the primary knowledge-base employed by The Uncanny Dream Machine – namely, the dream narrative corpus – should lead the examiner to accurately assess the possible range of narrative scenarios. At the level of plot, The Uncanny Dream Machine should thus be very predictable. Story coherence is largely preprogrammed first as a result of the style of the existing corpus texts, and second by the human-generated recursive transition network templates. This is, admittedly, the system's greatest weakness. Any new plots or scenarios cannot be generated computationally, but must instead be either added as new RTNs, or through additions or subtractions from the text corpus.

Linguistically, however, the system should produce significant variety in its texts. Such variety is largely due to the application of WordNet as a means of expanding the system's range of available characters, concepts, objects and descriptive vocabulary. As one of the largest lexical databases of the English language on the planet, WordNet lends the system significant linguistic expressivity. The downside of relying heavily on WordNet for linguistic expressivity is that it provides the system with too broad a vocabulary. The system's narratives occasionally use words not commonly found in everyday speech. This is particularly true of its use of adjectives. For instance, rather than use the common phrase “white socks,” the system may instead use any number of obscure synonyms for white, including “alabaster,” “achromatic” and “immaculate.” What results are sentences that are completely valid and sensical, but difficult for most listeners to comprehend.

Thus, while The Uncanny Dream Machine fails in its mission to induce déjà vu or regularly produce comprehensible texts, it succeeds in its ability to capture attention, engage an audience, and provoke conversation. Where it succeeds, it succeeds as provocative philosophical talking piece. Where it fails, it fails in aspects of its technical implementation. As an initial foray into the art of computationally generated dream-like narrative, this is to be expected. There is still much work to be done.

Chapter 6: Conclusion

The practice of expressive AI encourages cultural practitioners to apply the rich technical practices of traditional AI in new, meaningful, and non-traditional ways. These new applications enrich the field of AI, but also question its foundational assumptions. If we cannot ever build really intelligent or really creative disembodied machines, at what point do we abandon the endeavor towards doing so? If creativity is not something objectively *out there*, but rather a subjective, embodied act of meaning making, at what point do we recognize the tasks of traditional computational creativity as futile? No non-living machine will ever be creative. No non-living machine will ever dream.

The construction of a dreaming machine thus serves no practical purpose other than as a philosophical investigation. It spawns far more questions than it answers. It solves no problems. Under traditional paradigms *it* is the futile act. Yet, the act of applying computation to this investigative space nevertheless provides an opportunity to examine the nightly phenomenon we call dreaming in a completely different light. This light forces us to acknowledge that dreaming is not just a mental activity, it is a bodily activity. Dreaming, like storytelling, contains meaning only in relation to our everyday lived, embodied experience. This is ultimately the message that The Uncanny Dream Machine offers to tell: that as a disembodied zombie storyteller, its dreams are nothing compared to ours.

Bibliography

[1] Agre, P. (1997) "Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI" eds. Bowker, G., Gasser, L., Star, L. and B. Turner, *Bridging the Great Divide: Social Science, Technical Systems, and Cooperative Work*, Lawrence Erlbaum Publishers, Mahwah, NJ.

[2] Arduino Project (2010) *Arduino*, http://www.arduino.cc.

[3] Bird, S. (2010) *Natural Language Toolkit*, http://www.nltk.org/.

[4] Boden, M.A. (2004) *The Creative Mind: Myths and mechanisms*, Routledge, New York, NY.

[5] Brill, E. (1992) "A Simple Rule-Based Part of Speech Tagger" *Proceedings of the Third Conference on Applied Computational Linguistics*, Trento, Italy.

[6] Bringsjord, S. and D.A. Ferrucci (2000) *Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine*, Lawrence Erlbaum Associates, Mahwah, NJ.

[7] Brooks, R.A. (1991) "Intelligence without representation" *Artificial Intelligence*, 47:139-159.

[8] Bulhak, A. (1996) "On the Simulation of Postmodernism and Mental Debility using Recursive Transition Networks" Monash University Department of Computer Science.

[9] Cepstral LLC (2010) *Cepstral Text-to-Speech*, http://cepstral.com/.

[10] Colombetti, G. and E. Thompson (2009) "The Feeling Body: Toward an Enactive Approach to Emotion" eds. W.F. Overton, U. Mueller and J. Newman, *Body in Mind, Mind in Body: Developmental Perspectives on Embodiment and Consciousness*, Erlbaum.

[11] Cycorp, Inc. (2010) *Cyc*, http://www.cyc.com/.

[12] Dick, P.K. (1996) *Do Androids Dream of Electric Sheep?*, Del Rey, New York, NY.

[13] Domike, S. (2003) "The Recombinant History Apparatus Presents: Terminal Time" eds. Mateas, M. and P. Sengers, *Narrative Intelligence*, John Benjamins.

[14] Ekbia, H.R. (2009) *Artificial Dreams*, Cambridge University Press, New York, NY.

[15] Freud, S. (1919) "Das Unheimliche" *Imago*, 5-6:297-324.

[16] Freud, S. (1967) *The Interpretation of Dreams*, Avon Books, New York, NY.

[17] Gallese, V. and G. Lakoff (2005) "The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge" *Cognitive Psychology*, 21:1-25.

[18] Gelfand, B., Wulfekuhler, M. and W. Punch III (1998) "Automated Concept Extraction from Plain Text" *Papers from the AAAI 1998 Workshop on Text Categorization*, Madison, Wisconsin.

[19] Hanna, R. and E. Thompson (2003) "The Mind-Body-Body Problem" *International Journal for Interdisciplinary Studies*, 7:23-42.

[20] Hanson, D.H. (2010) *Ruby on Rails*, http://rubyonrails.org/.

[21] Harnad, S. (1990) "The Symbol Grounding Problem" *Physica D*, 42:335-346.

[22] Harnad, S. (1995) "Why and How We are Not Zombies" *Journal of Consciousness Studies*, 1:164-167.

[23] Hofstadter, D. (1999) *Gödel, Escher, Bach: an eternal golden braid*, Basic Books.

[24] Jentsch, E. (1906) "Zur Psychologie des Unheimlichen" *Psychiatrisch-Neurologische Wochenschrift*, 8.22:195-98.

[25] Kafka, F. (2008) *The Metamorphosis*, Aventura Press.

[26] Kinoshita, M., Shimada, N. and K. Arikawa (1999) "Colour Vision of the Foraging Swallowtail Butterfly Papilio Xuthus" *The Journal of Experimental Biology*, 202:95-102.

[27] Lemmen, R. (1996) "Embodiment for Non-Cartesians" University of Sussex at Brighton.

[28] Llinás, R.R. and D. Paré (1991) "Of Dreaming and Wakefulness" *Neuroscience*, 44:521-35.

[29] Markram, H. (2006) "The blue brain project" *Neuroscience*, 7:153-60.

[30] Mateas, M. (2001) "Expressive AI: A hybrid art and science practice" *Leonardo*, 34:147-153.

[31] Mateas, M. and P. Sengers (1998) "Narrative Intelligence" *Proceedings of the 1998 AAAI Fall Symposium*, Orlando, Florida.

[32] Maturana, H.R. and F. Varela (1987) *The tree of knowledge: The biological roots of human understanding*, Shambhala Publications, Boston, MA.

[33] Meehan, J. (1976) *The Metanovel: Writing Stories by Computer*, Dissertation, Yale University.

[34] Miller, G. (2010) *WordNet*, http://wordnet.princeton.edu/.

[35] Mueller, E. (1990) *Daydreaming in Humans and Machines*, Ablex Publishing, Norwood, NJ.

[36] Murray, J. (1998) *Hamlet on the Holodeck*, MIT Press, Cambridge, MA.

[37] Nagel, T. (1974) "What Is It Like to Be a Bat?" *Philosophical Review*, 83:435-450.

[38] Nass, C. and L. Gong (2000) "Speech Interfaces from an Evolutionary Perspective" *Communications of the ACM*, 43:36-43.

[39] Nass, C., Steuer, J., Henriksen, L. and D. C. Dryer (1994) "Machines, social attributions and ethopoeia: performance assessments of computers subsequent to 'self-' or 'other-' evaluations" *Int. J. Human-Computer Studies*, 40:543-559.

[40] Nelson, D.L., Mc Evoy, C.L. & T.A. Schreiber (1998) *The University of South Florida word association, rhyme, and word fragment norms*, http://www.usf.edu/FreeAssociation.

[41] OpenWrt (2010) *OpenWrt*, http://openwrt.org/.

[42] Penny, S. (1995) *Petit Mal: an Autonomous Robotic Artwork*, http://ace.uci.edu/penny/works/petitmal.html.

[43] Pérez y Pérez, R. and M. Sharples (2004) "Three Computer-Based Models of Storytelling: BRUTUS, MINSTREL and MEXICA" *Knowledge-Based Systems*, 17:15-29.

[44] Piaget, J. (1951) *Play, Dreams, and Imitation in Childhood*, William Henemann, London.

[45] Plutchik, R. (1994) *Emotions and life: perspectives from psychology, biology, and evolution*, American Psychological Association, Washington, DC.

[46] Price, R. (1978) *A Palpable God: Thirty Stories Translated from the Bible*, Anthenum, New York, NY.

[47] Revonsuo, A. (2000) "The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming" *Behavioral and Brain Sciences*, 23:793-1121.

[48] Rokeby, D. (1998) *The Giver of Names*, http://homepage.mac.com/davidrokeby/gon.html.

[49] Searle, J. (1980) "Minds, Brains and Programs" *The Behavioral and Brain Sciences*, 3:417-457.

[50] Searle, J. (1990) "Is the Brain a Digital Computer?" *Proceedings of the American Philosophical Association*, 64:21-37.

[51] Searle, J. (1992) *The Rediscovery of the Mind*, MIT Press, Cambridge, MA.

[52] States, B. (1993) *Dreaming and Storytelling*, Cornell University Press, Ithaca, NY.

[53] Thompson, E. and D. Cosmelli (2009) "Embodiment or Envatment? Reflections on the Bodily Basis of Consciousness" eds. Gapenne, O. and E. Di Paolo, *Enaction: Towards a New Paradigm For Cognitive Science*, MIT Press, Cambridge, MA.

[54] Turing, A. (1950) "Computing Machinery and Intelligence" *Mind*, LIX:433-460.

[55] Turner, S. (1994) *The Creative Process: A Computer Model of Storytelling and Creativity*, Lawrence Earlbaum Publishers, Hillsdale, NJ.

[56] Varela, F., Thompson, E. and E. Rosch (1993) *The Embodied Mind*, MIT Press, Cambridge, MA.

[57] Von Uexküll, J. (1957) "A Stroll Through the Worlds of Animals and Men: A Picture Book of Invisible Worlds" eds. Schiller, C.H., *Instinctive Behavior: The Development of a Modern Concept*, International Universities Press, New York, NY.

[58] Wardrip-Fruin, N. (2009) *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*, MIT Press, Cambridge, MA.

[59] Weizenbaum, J. (1976) *Computer Power and Human Reason*, W.H. Freeman, San Francisco, CA.

[60] Wittgenstein, L. (2001) *Philosophical Investigations*, Oxford, Malden, MA.

[61] Zhu, J. and D.F. Harrell (2009) "System Intentionality and the Artificial Intelligence Hermeneutic Network: the Role of Intentional Vocabulary" *Proceedings of the 2009 Digital Art and Culture Conference*, University of California, Irvine.

Appendix A: Sample Dream Narratives

Sample 1: Anger

I am in a garage with a girl and some people. We are in the dressing room of the garage, just hanging out. The room is like a room at my parents' house. Rather suddenly, some genius comes and tries to rob me. I get sore and scold the genius. The genius pulls out a firearm and points it at a girl. I tackle the genius, hitting the ground.

Sample 2: Embarrassment

I am at my college in front of a large gathering. I look down and find that I'm taking a shower. The gathering starts to laugh at me. What can I do? I see no exit. An actor appears and tells me to go to my dormitory room. I say, “I know!”

Sample 3: Anxiety

I am waiting on a chair in a foreign island. It is loud; many chatting people. Apparently, I'm here for a comprehensive examination, but I don't remember why. A native fisherman appears and tells me to leave before its too late. Before I know it, Benjamin Linus traps me and puts me in an operating room. I start to panic. Benjamin Linus laughs. Benjamin Linus injects me with methadone.

Sample 4: Joy

I seem to be playing a percussion instrument in a big band called The Cars. I don't know how to play, but the rock band doesn't seem to mind. Clapping to the music, I start to play an electronic instrument. In the audience I can see my uncle. I delight and delight again.