A Theoretical Cognitive Construct of a 3D Embodied Agent:
VAL, the Virtual Autonomous Learner

Lexx L. Lazerman

A Thesis

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Educational Technology) at
Concordia University
Montreal, Quebec, Canada

April 2012

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: **Lexx L. Lazerman**

Entitled: **A Theoretical Cognitive Construct of a 3D Embodied Agent: VAL, the Virtual Autonomous Learner**

and submitted in partial fulfillment of the requirements for the degree of

Master of Arts (Educational Technology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Frederica Martin	Chair
Dr. Steven Shaw	Examiner
Dr. Robert Bernard	Examiner
Dr. Vivek Venkatesh	Supervisor

Approved by Dr. Richard Schmid
Chair of Department or Graduate Program Director

Dr. Brian Lewis
Dean of Faculty

March 30, 2012

Abstract

A Theoretical Cognitive Construct of a 3D Embodied Agent:
VAL, the Virtual Autonomous Learner

Lexx Lazerman

The cognitive sciences have always educated educators by providing a pedagogical framework as a guide. However, the standard cognitive sciences are being challenged by a new paradigm, embodied cognition, in which learning is part of a dynamical system. In this paradigm, virtual embodiment (VE) is the new artificial intelligence (AI). This thesis is an application of VE, introducing an approach to developing a virtual 3D agent that has the potential to achieve "strong AI" status. I believe such agents can mature into AI educators. And that the development of a *great* AI educator starts with the development of a humble AI child. My methodological approach is a metasynthesis of a broad range of disciplines and consists of (1) the use of *empirical research* to ground my ideas, (2) the *integration of dissimilar research* to construct new ideas, and (3) the use of *thought experiments* to uncover the fundamental nature of learning within an embodiment paradigm. As a result, this thesis introduces a virtual 3D agent, the virtual autonomous learner (VAL), along with key elements of its ecological construct. With an embodied cognitive perspective, VAL seeks to find its own affordance and that of its environment. I conclude that (1) the construct for VAL needs to accommodate different cognitive architectures if we are to make full use of its methodology; (2) a rigorous virtual curriculum must be developed, and efficient pedagogical tools should be designed and developed to implement this curriculum; and (3) an educational perspective is paramount for this project.

Acknowledgements

I would like to thank my supervisor, Dr. Vivek Venkatesh for his support and professionalism. I would like to thank my committee members Dr. Steven Shaw and Dr. Robert Bernard for their comments which have given me much to think about for years to come. And I would like to thank a former committee member, the late Dr. Gary Boyd for his inspiration. I would also like to thank friends and colleagues, like Konstantina Michos, Terry Lambert, Jordan Lallouz, MinYing Huang and Wilhelmina Wong for their support and believing in me. Finally, I would also like to thank Dr. Ginger Campbell and her “Brain Science Podcast” which provided me with many hours of stimulating scientific conversation. A special thanks to Kara Farra, who paid a dear price for this thesis since it stole much of our valuable time together.

Table of Contents

List of Figures

Introduction

The cognitive sciences have always educated educators on learning. The three main theoretical schools, behaviorism, cognitivism, and constructivism, provide a pedagogical framework to guide educators. However, the standard cognitive sciences are being challenged by a new paradigm, embodied cognition, in which learning is part of a dynamical system (Shapiro, 2011). This thesis is an application of the new paradigm. Insights about embodied cognition are transforming traditional cognitive science, and virtual embodiment (VE) is the new paradigm for artificial intelligence (AI).

AI is everywhere and used by everyone. If you recently Googled something—anything—then you recently used AI. Broadly speaking, AI includes any algorithm that learns. AI can take on many forms as it is applied in different domains. When it is used in education, most AI resides in intelligent tutoring systems (Beck, 1996). In essence, these systems are "AI educators." This thesis takes the concept of the AI educator to a new level.

AI includes a wide variety of processes that simulate intelligent behavior. AI systems can be subdivided into two classes based on their functionality: (1) weak AI and (2) strong AI (Searle, 1980). Weak AI systems are systems that do not match or exceed human intelligence; these are software systems such as pattern recognition systems, expert systems, and data mining systems. Today these weak AI systems are used in a wide range of fields, including medical diagnosis, stock trading, robotic control, gaming, and the toy industry. Most relevantly for our purposes, weak AI is being used in education. In contrast, unlike weak AI systems, strong AI systems that match or exceed human intelligence have not been realized.

In this thesis, I will explore the potentials of VE and begin a discourse on the construction of a virtually embodied synthetic learner that serves as a viable candidate for achieving strong AI. Most of this thesis centers around the creation of a construct that contains a 3D virtual agent; I argue that embodiment is a viable concept that can aid in the development of strong AI. Although this thesis presents the model for a 3D virtual computer agent, the Virtual Autonomous Learner (VAL), it is important to note that the creation of VAL is just the first step in the development of strong AI. Following the creation of VAL, the educational technology development device called "rapid prototyping" will be applied to accelerate learning. Rapid prototyping is used in many design-related domains and will be applied to VAL's curriculum, with the ultimate goal of having VAL achieve strong AI status. The end product of such an achievement would be an AI educator.

Key features of VAL's construct are described and discussed throughout this thesis, including the accommodation of different cognitive architectures, the development of a rigorous virtual pedagogy, and the adoption of an educational perspective. From an embodiment perspective, these are all key elements for the development of strong AI.

This thesis contains six sections. In the first section following this introduction, "Educational Rationale and Context," I discuss some specific shortcomings of human educators and argue that there is a need for AI educators. The next section, "Theoretical and Conceptual Framework," covers the angle from which I approach my topic; here I identify the assumptions and presuppositions that I bring to my research. The section "Literature Review" contains a broad review that provides the background information needed to understand important concepts. In this section, I discuss similar research and technologies, occasionally noting how the research has influenced my work and often commenting on how other technologies fall short of my research goals. In the following section, "Methodology," I articulate the process that I used to generate my ideas.

The final two sections, "An Ecological Construct of a Virtual Autonomous Learner" and "Learning Methodology of a Virtual Autonomous Learner," are the main sections of this thesis. In the former, I present the key elements that are needed for an embodied agent to obtain strong AI status, and in the latter, I lay out the advantages of learning in a virtual environment rather than the real world. I conclude this thesis with a discussion of what can be expected from VAL and suggestions about future research directions.

Educational Rationale and Context

Educators have the most profound jobs, period. After all, much of civilization arose from individuals teaching individuals. In the beginning, the human educational system was explicit and relied heavily on oral tradition. Then, around eight thousand years ago, humans developed writing systems that allowed off-loading some of the knowledge and information that was stored in the heads of educators. Around five hundred years ago, the printing press made information widely available to anyone who could read. Today, information and knowledge are undergoing another revolution of sorts, a revolution whose poster child is Google. Google's mission statement, "Organize the world's information and make it universally accessible and useful" (Google, 2012), is not a collection of idle words—Google is in the process of digitizing every book ever written! Significantly, today there are more and more sophisticated devices (computers) that off-load more and more information from educators. The significance of this revolution with regard to education is that it has given students unprecedented access to knowledge and information. Yet we still have teachers. Why? Because, even though information is more readily available than ever before, teachers are still needed to guide inquisitive minds. As this thesis will show, this is especially true in early education where much learning comes from the interactions between teachers and students. That is, even in light of the digital revolution, face-to-face communication remains an efficient way to exchange knowledge and ideas.

Evidently, then, the human educator still has much value. Nevertheless, humans are fundamentally flawed as educators. One of these flaws is the simple fact that all humans—educators included—have biases; this is just part of our nature. It is also widely accepted that the influence an educator can have on students is profound and that the influence can be for the better or for the worse. For example, studies have shown that teachers' expectations of students can influence performance (Good, 1981).

Knowing that every educator, like every person generally, has biases, what can be done to ensure that students are not negatively affected? The answer is that not much can be done. Even the best educators, those who acknowledge their biases and try not to "act" on them, cannot control all of their biases, because many of these are subconscious (Gladwell, 2005). It would appear that the quality of an educator is linked to the quality of the individual. With that said, even though we know that some educators can potentially have a negative influence on students, we still value face-to-face or student-to-teacher interactions.

Furthermore, even the best educators—those who are considered among their peers as exceptional educators and who follow "best practices," modifying their pedagogical approaches for each student with a recognition of the student's strengths and weaknesses and giving each student what that student needs when it is needed—have limitations. A measurable limitation of educators is inefficiency. There are only so many hours in the day, and there are many, many students. In fact, I would argue that if educators are tools for educating students, then it would be logical to surmise that a human educator, like any other tool, can be replaced by more efficient computer technology. Unlike a computer, a human educator cannot copy itself, is not portable, cannot update the latest "best practices," and is not accessible in remote corners of this planet in the way that an AI educator can be.

Is it possible to have the best of both worlds? Can we create something that looks and acts like a human educator, yet does not have human biases and *does* have the efficiency of a computer?

I ask you to suspend your disbelief for a moment and imagine a perfect educator. What would your perfect educator look like? How would that educator behave? The big question is: how would *you* design a perfect educator? This is perhaps a futile question, for if we asked a thousand educators for their vision of a perfect educator, we would surely get one thousand different visions. Regardless of what *your* idea of the perfect educator might encompass, consider this: one thing all educators have in common is that they each had an education themselves. This may seem like a trivial fact, but it is not—it is a profound one. It should not be taken lightly, because it is often life experiences that make an ordinary educator a great educator. So, regardless of how you want to design your perfect educator, that educator will have to be a student first and will have to develop, learn, make meaning, and make sense of the educator's environment before knowledge can be shared. This thesis argues that the development of a great AI educator starts with the development of a humble AI child. The thesis starts with the belief that humanity needs more and better educators, journeys through an exploration of the new AI paradigm, VE, and ends with a virtual child that will one day educate us.

As we have mentioned, AI can provide a nonbiased approach to instruction, and it is already used in education in the form of intelligent tutoring systems. Indeed, in the past decade or so, virtual environments have started to gain momentum in educational settings. However, AI and virtual environments each have shortcomings with regard to their effectiveness in educational settings. Therefore, this thesis explores the potential of combining AI and virtual environments by applying embodied cognition principles to a virtual agent. Through this exercise, we will learn much about the nature of learning.

> Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's?
>
> (Turing, 1950, p. 23)

Looking at the history of AI, it is clear that Alan Turing understood the profound potential of this emerging realm. He proposed two approaches to developing intelligent machines: (1) creating an adult mind and (2) creating a child (Turing, 1950). Thus, the idea of developing a synthetic child, a child "that would learn from experience just as a human child does" (McCarthy, 2008, p. 1), is *not* new. What *is* new since Turing's time is our greater understanding of cognition, our ability to study brain activity in real time, and, most relevant to my thesis, our technological proficiency for creating virtual environments.

The traditional cognitive sciences, with their symbolic computations, had a strong influence on traditional AI (weak AI). Since this thesis views embodied cognition as the foundation for the new paradigm of cognitive science, VE should therefore become the new AI (strong AI). With that said, I shall argue—and I hope that you will agree after reading this thesis—that creating a virtual environment and placing a virtual infant in this environment is an efficient approach to developing humanlike intelligence.

Along with advocating the virtual environment approach, this thesis also argues that some of the issues relating to the design and development of intelligent machines are inherently "nature vs. nurture" problems. That is, there are two issues that contribute to the development of intelligent behavior: (1) on the nature side, there is the issue of the metaphoric "blank slate," and (2) on the nurture side, there is the issue of how one would educate a virtual child (McCarthy, 2008).

To complete my presentation of the educational rationale and context for this thesis, I need to mention one caveat concerning the proverbial "blank slate." It is important to note that this is borrowed terminology, and I use it as a conceptual foundation because of its colloquial familiarity. However, the term is not accurate because the slate is never completely blank. That is, it is widely accepted that living DNA stores innate information, although how and how much is a matter of great debate. Nevertheless, in this thesis, the metaphoric blank slate is more like an ever-changing template, a cognitive template. Therefore, I use the term "template" rather than "blank slate" to refer to the default configuration of the virtual agent's cognitive architecture.

This thesis provides an integrative review of many domains to illustrate the potentials of 3D agents, and it introduces VAL, the Virtual Autonomous Learner, as a candidate for achieving strong AI.

Theoretical and Conceptual Framework

The theoretical framework for my thesis has its foundation in many domains, including computer science, neuroscience, psychology, philosophy, evolution, and education.

From a philosophical and practical point of view, this thesis rejects dualism and takes as its theoretical foundation the belief that thoughts and behavior are the result of neural activity. This thesis takes a physical reductionist point of view, holding that physical phenomena can be explained by breaking large complex elements down into smaller simple parts. And since, as just mentioned, behavior is generated from neural activity, it should be possible to generate behavior from artificial neurons in a virtual environment. This thesis takes an embodied cognition point of view with regard to learning, as opposed to the standard model of the cognitive sciences that focuses on symbolic computation. That is, much of the conceptual framework for this thesis is based on the belief that intelligence emerges from the dynamic interaction of the body, the mind, and the environment. Therefore, the framework for this thesis rejects the old AI paradigm as a viable means for achieving strong AI and incorporates the belief that any AI system that tries to generate humanlike intelligence needs a humanlike body.

Further, this thesis looks to evolution for insight into how the mind works, postulating that an examination of evolutionary stages can reveal hidden secrets about how the mind works.

Finally, this thesis uses Moore's Law to partially justify its potential success, to refute old AI paradigms, and to argue that virtual environments can exploit principles of embodied cognition.

Literature Review

The interdisciplinary nature of strong AI calls for a broad review of the literature, and this review, along with the material in the section "Theoretical and Conceptual Framework," should help readers understand the perspective that underlies this thesis.

Generally speaking, three approaches have been taken in the development of AI: the symbolic, statistical, and embodied approaches. Symbolic processing has been the mainstream approach to the development and application of AI systems. Although the symbolic approach has seen some limited success, it does not have the necessary conceptual framework for achieving the status of strong AI. Statistical learning algorithms, in conjunction with artificial neural networks (ANNs), are relatively more successful and are widely used in medical diagnosis, stock trading, robot control, gaming, and the toy industry. However, it is unlikely that an autonomous learner can be developed through symbolic methods, mainly because these methods lack the necessary ingredients for autonomous learners, which include having a body. The relatively new embodied approach to AI seeks to develop intelligence through the interaction between an agent and its environment. Researchers have been using the embodied approach to integrate human features such as hearing, vision, and emotions into virtual bodies (Lorenz & Barnard, 2007).

Hence, a literature review for the application of embodied cognition principles to a virtual autonomous 3D agent with the purpose of developing strong AI must cover research in three main domains: the environment, the body, and the mind. To better inform the reader about these three domains, this literature review tries to answer three questions:

1. What is the current state of technology with regard to virtual environments?
2. What is the current state of technology with regard to autonomous agents?
3. What would the mind of an autonomous agent look like?

Real and Virtual Environments

A wide range of virtual environments are available for researchers to explore embodied cognition principles. Some environments are simple and resemble the popular website SecondLife, while others are very sophisticated and contain dynamic elements that can simulate real world physics (Boulos, Hetherington, & Wheeler, 2007).

Some of the simple virtual environments have been used effectively in educational settings and offer a rich method of multimodal interaction between 3D agents and humans. To explore the effectiveness of virtual agents in educational settings, Rickel & Johnson (2000) designed the 3D agent Steve (short for "Soar Training Expert for Virtual Environments"). Steve can communicate with students through voice recognition and a voice synthesizer. Steve also uses gaze, gesture, and body orientation to guide the student in a particular task, in much the same way that a real educator would. As the authors put it, "Steve illustrates the enormous potential for face-to-face, task oriented collaboration between students and synthetic agents in virtual environments" (Rickel & Johnson, 2008, p. 24). However, although some elements of Steve's construction can be considered to be AI, most are not. In its environment, Steve's movement (behavior) is preanimated: it can only move in a predetermined way (Rickel & Johnson, 2000). Having said that, however, I'll add that many elements of Steve's environment *are* similar to what I discuss in my thesis, and the application of Steve as a virtual instructor is a good model for thinking about the end use for VAL.

Other virtual systems and environments are strictly centered around and based upon artificial neural network developmental tools. These environments allow researchers to explore the complexities of cognitive neuroscience in virtual 3D environments. For example, Emergent, an open-source tool, demonstrates how artificial neural networks can be used to simulate large-scale cognitive phenomena (O'Reilly & Munakata, 2000). This software system incorporates anatomical and physiological properties of the neocortex, basically simulating neurons down to the neural-transmitter level. This tool makes it possible to simulate large-scale brain networks with memory systems, sensors capable of pattern recognition, and artificial neurons capable of motor control. The relevance of this tool is that the same neural simulation technology can be used as the backbone for the virtual autonomous learner's control system, as we shall see in the section "An Ecological Construct for the Virtual Autonomous Learner" of this thesis.

There are more encompassing AI environments that contain humanlike agents that bridge the gap between the virtual and real worlds. For example, SIMNOS, a virtual embodied agent developed at The Machine Consciousness Lab at the University of Bristol, simulates the cognitive development of an agent within the virtual world (Gamez, 2008). After completing a virtual learning intervention, the system can upload the "mind" of the agent to its real world counterpart, the physical robot CORNOS. That is, the cognitive architecture of this system can be used in both the real world and the virtual world because SIMNOS's virtual representation and control system are immersed in an environment with a "real time physics-based" emulator (Gamez, 2008, p. 895).

Even with a physics emulator, there is a stark contrast between virtual and physical environments in most respects. However, many researchers believe that the real and virtual environments pose "nearly identical challenges from the agent modeling perspective" because in both domains, agents receive sensory data from the environment (Best & Lebiere, 2006, p. 186). Furthermore, today's virtual worlds can simulate most real-world elements; we can model and simulate the physical world in minute detail down to the ion channels of neurons. Moreover, within these virtual worlds, we can simulate the real-world physics that is necessary for an embodied approach to learning. For example, dynamic physics engines can simulate forces like gravity. Virtual environments can also simulate surface textures that affect their virtual properties, like friction. Also within these environments, optical properties like reflection and refraction can be simulated.

Virtual environments are so apt at representing real-world physics and dynamics that engineers use data generated in virtual environments to guide the construction of real-world artifacts. All of this means that a methodological approach that uses virtual environments allows researchers to study the relationship between the environment, the body, and the mind. Virtual worlds are modeled on real-world systems so that the stimulus data for an agent's system are essentially the same as real-world data (Best & Lebiere, 2006). Best and Lebiere add that "most importantly, these domains involve agents that [can] interact with humans and each other in real-time in a three-dimensional space" (2006, p. 186).

Embodiment

As mentioned in the section "Educational Rationale and Context," this thesis has a bias towards Turing's (1950) child-based approach to the development of strong AI, believing that it is technologically feasible to build "a child machine that would learn from experience just as a human child does" (McCarthy, 2008, p. 2003). McCarthy believes that the past failure of AI to create intelligence by creating a child owes to the fact that "innate knowledge the child machine should be equipped with was ignored" (McCarthy, 2008, p. 2003). I shall further comment on this issue of innate knowledge in the section "Learning Methodology for a Virtual Autonomous Learner."

As the literature illustrates, the idea of a synthetic child learner has been around since the inception of AI. However, the tools to implement this idea have only been available to researchers in recent years. This is important, because much of what I present not only uses embodied cognition, the new paradigm for the cognitive sciences, but it also takes advantage of these new software tools. Research that is similar to mine, both in the physical world and in a virtual world, also takes advantage of new technologies and techniques. Two examples follow.

The physical robot iCub is an open-source humanoid robot that has the approximate size and proportions of a human baby. This is an autonomous agent that contains an array of sensors and motors that simulate human faculties. The researchers who developed iCub claim that it has "the ability to learn the affordances of objects" (Metta, 2010, p.7). The cognitive architecture of this robot enables learning and development in a social environment; that is, it learns from its interactions with people.

The researchers also claim that iCub "adapts and learns how to behave in new situations" and "invents new solutions on the basis of the past experience" (Metta, 2010, p. 1).

In the virtual realm, the CyberChild may be the most ambitious project that attempts to simulate the enormous complexity of the human body, and this project is most similar to my own research. CyberChild is a virtual child agent that includes features like hearing and touch receptors, pain receptors, muscles, and a gut with bladder control (Cotterill, 2003). Cotterill believes that virtually modeling the core elements of the mammalian nervous system that represent "the earliest stages of human life" will bring about intelligent behavior (2003, p. 31). However, Cotterill's project focuses on the emergence of consciousness, and in this respect, our research diverges. Although our research methods are similar in some respects, my work focuses on producing a virtual agent that learns autonomously.

Other researchers, like Mueller & Minnery (2008) and Goertzel & Pennachin (2007), adopt a similar embodiment perspective. However, much of their research focuses on high-level cognitive processes and resembles the inquiry methods of the old AI paradigm, whereas CyberChild and iCub are examples of recent research projects that adopt the new paradigm of looking at child embodiment for insight about intelligence.

Generally speaking, embodiment as a research field has been around for at least 20 years, and Rodney Brooks may be one of its most prominent advocates. Brooks has written extensively on robotics systems that depart from classical AI assumptions. Brooks discusses how to generate intelligent behavior in robots by exploiting the morphology and the physical properties of the body. He advocates embodied cognition principles that promote learning, a key component of my thesis. Brooks develops perceptual systems that incorporate visual, vestibular, auditory, tactile, and kinesthetic senses to create what he calls "a complete embodied system" (Brooks, 1999, p. 52). Pfeifer, Bongard, and Grand (2007) use the term "complete-agent principle" (p. 104) and apply many of Brooks' ideas, offering detailed examples and empirical studies of applications of embodied cognition principles to agents.

Brain and Mind

While the previously mentioned researchers concentrate on the morphology of the body and the relationship between the body and the environment, other researchers offer insight into the morphology of neurons. To shed light on the question "What would the mind of an autonomous agent look like?" we must review the morphology of neurons (their architecture) and the morphology's effects—cognition. Knowledge of how human neural morphology changes over time and how that change affects overall behavior can guide the development of an artificial neural network for a complete agent.

Different research domains can be combined to paint a clearer picture of the brain–mind relationship. For example, Edelman and Mountcastle's (1978) study of higher brain functions offers insight into the morphology of the brain's neocortex, and Gopnik's (2009) theory of babies' consciousness offers insight into the mind. Edelman "Neural Darwinism" proposes that neurons are first indiscriminately and massively interconnected during early stages of development and that the neural network is subsequently pruned through experience, keeping only the strongest connections. This theory complements Gopnik's ideas about babies' consciousness: Gopnik sheds light on the mind, while Edelman and Mountcastle shed light on the brain. These authors would lead us to believe that babies are more aware (conscious) of and less biased about their environment than are adults. The section "Learning Methodology for a Virtual Autonomous Learner," will make it evident why this is significant.

Gopnik (2009) also claims that infants are logical and rational and understand cause and effect, somehow making complicated calculations with conditional probabilities. Apparently, a baby is the "most powerful learning computer on the planet" (Gopnik, 2011) and uses Bayesian learning algorithms.

Generally, solid insights about the human mind can contribute to the development of strong AI. In this thesis, the mind encompasses all cognitive functions, and from an embodied cognitive point of view, cognitive functions incorporate both the brain and body.

Any serious discussion of the mind or humanlike intelligence must account for consciousness. There are many views about the mind and human consciousness that can be modeled in the pursuit of strong AI. Therefore, when they conduct research and develop theories related to the mind, many researchers look for evolutionary factors that might have contributed to human intelligence (Greenspan & Shanker, 2004; Gopnik, 2009; Hawkins & Blakeslee, 2004; Jaynes, 1990; Montague, 2006). In fact, I believe that metaphorically rewinding the stages of human evolution can offer the most valuable insight into the development of embodied agents' cognitive architectures.

Hawkins and Blakeslee's (2004) theory of intelligence offers a starting point for a simple cognitive architecture. These authors offer a plausible explanation of the major function of the human neocortex. They interpret Mountcastle's (1997) cortical column as a mini generic processor, and they claim that the neocortex is a massive memory-prediction machine. The sample generic architecture that I offer in the section

"Subsystem" of "Learning Methodology for a Virtual Autonomous Learner" is based on Hawkins and Blakeslee (2004).

Greenspan and Shanker (2004) lay out a developmental road map of human behavior and intelligence. These authors believe that higher cognitive functions like abstract thinking began with basic emotions that successively became more and more complex. This view fits well with an embodied approach to strong AI. An application of their ideas would be most fruitful as a guide for the development of VAL's curriculum. Jaynes (1990) suggests a plausible method for developing a virtual embodied conscious mind, and this might contribute greatly to the general architecture of a complete virtual agent. Jaynes' theory, which is that the conscious mind has only been around for a few thousand years, provides insight into the type of cognitive performance that might be expected from an embodied agent. Jaynes' theory is the reason that I shy away from claiming that VAL will develop a conscious mind at any time in the near future. My reasoning is that if the conscious mind has only been around for a small period of human history, then this may mean that consciousness appears at a very high level of cognitive development—and therefore, there is no need to look for consciousness in VAL, especially during the early stages of development. This does not mean that we should not strive for developing consciousness in virtual agents; in fact, Jaynes offers an interesting model of how an agent can acquire a conscious mind. The bicameral mind is a model of the mind in which the mind is slightly divided (this can be seen physically with the corpus callosum) and one side can become aware of the other. This theory about the mind can be modeled with artificial neural networks.

This application of Jaynes' (1990) research was not his intent. That is, his book does not discuss computation or artificial neural networks, so I have taken great liberty in translating his ideas and applying them to my research. However, other researchers with similar ideas *do* have more computational backgrounds. A good example is Reed Montague, a professor of neuroscience. Montague's (2006) discussions and ideas center around computation, efficiency, and value, and he discusses a particular goal common to all animals: the goal of surviving. Montague discusses how our brains are like computers and postulates that our brains have evolved to survive by seeking simple goals like food, water, and sex. He claims that these simple goals can then turn into more complex ideas. Most importantly, he discusses how the brain builds meaning through a built-in value system. The relevance to my thesis is that Montague offers insight into computational perception and answers the very important question, "How would a complete agent build meaning?"

These theories of mind and consciousness that pertain to the architecture of neural connectivity are very important in the development of strong AI. As already mentioned, past failures in strong AI have been attributed to the lack of innate attributes within the architecture of artificial neural networks (McCarthy, 2008). With the proper innate neural configurations and exposure to a stimulating environment, people thrive (Ridley, 2003), and I believe that the same will be true of VE agents.

The Big Picture

An embodied cognitive perspective on developing strong AI seeks to find an algorithm for the relationship between the environment, the body, and the mind. Gibson (1979) articulates this relationship in what is popularly known as "affordance theory."

Gibson's theory discusses the surface and texture of objects, and he claims that how an animal perceives an object will determine how the animal will use it. Gibson holds that the morphology of an animal affects its perception. Gibson's discussion of affordances, the properties of objects that permit specific actions to be taken, is an important contribution to the overall theory of embodied cognition and therefore will serve as a good starting point for my discussion of which innate properties should be programmed into an agent. In many respects, my research is an application of Gibson's work. I shall say more about innatism in the section "Learning Methodology for a Virtual Learner."

The theory of affordance also offers indirect insight into how artificial vision can be developed. Gibson's views on perception are discussed and applied in Marr (1982), which provides ideas about how to develop better perception for virtual agents. Marr claims that vision is a complex information-processing system, and he deconstructs vision into many components that can be used as a computational filtering system. This area of research is very important because one of the most profound challenges facing the development of strong AI is the development of perception. Marr's deconstruction of vision complements Gibson's environment/body/mind processes. In his research, Marr describes how the brain processes different visual elements like edges, reflection, textures, and shadows. An application of this research, the creation of a multilayered neural network that filters different aspects of our visual world, inspired many AI researchers to research and create better visual perception devices for intelligent agents.

Today, we have the technology to exploit and apply some of the fundamental aspects of affordance theory. For example, most 3D modeling and animation software has the ability to create objects with surface textures that contain all of Marr's visual elements, providing 3D agents with the opportunity to learn the affordance of said objects.

Before I turn to the final topic in this literature review, there is one more important point that must be explicitly stated, as it is a testament to the state of the technologies discussed in this literature review: a synergy is developing in VE research, perhaps in part because many of the software packages and technologies that I have reviewed, like the dynamic physics simulators, are open source.

Finally, this literature review will take a quick look at the current state of technology in general and how it affects my thesis. The futurist Ray Kurzweil holds that the exponential growth of technology and its effect on society are profoundly shaping the new direction of AI. Kurzweil (1999) predicts that strong AI will be achieved by 2029, when he believes that computers will commonly pass the Turing test (Turing 1950). Others, most notably Paul Allen, the cofounder of Microsoft, believe that this will be achieved at a much later date (Allen & Greaves, 2011). Kurzweil (2005) bases much of his prediction on Moore's Law, which states that computer processing power will double, and the price per unit will drop by half, every 18 months. Moreover, this popular statement is only part of the big picture. That is, the price per unit is not the only factor that contributes to the exponential growth of computer technology; new algorithms that process information in new and more efficient ways also increase computational power. The success of strong AI is therefore not solely linked to processing power; innovative techniques will be the backbone for its success. It is new ideas and processes (algorithms) that will primarily contribute to the ultimate goal: a virtual agent that has strong AI status.

Methodology

My methodological approach—that is, the process by which I generate my arguments and my conclusions—is a metasynthesis of a broad range of disciplines and consists of (1) the use of *empirical research* to ground my ideas, (2) the *integration of dissimilar research* to construct new ideas, and (3) the use of *thought experiments* to uncover the fundamental nature of learning within an embodiment paradigm.

Empirical Research

Most ideas and theories need to be grounded on some starting point. Therefore, I humbly stand on the shoulders of other researchers and use their empirical research to ground my ideas. As you have seen in my literature review, the research that I use is broad and covers many disciplines. When reviewing a study or even a philosophical paper, I constantly ask myself, "How is this relevant to my research?"

Integrate Dissimilar Research

The cognitive sciences are by their very nature interdisciplinary; AI is only a small piece of a much larger picture that includes psychology, philosophy, neuroscience, linguistics, anthropology, sociology, and education. What might seem to be dissimilar areas of research are in fact intricately related. Moreover, the triangulation of data or knowledge from different domains paints a clearer picture of the very nature of VE. Therefore, crossing traditional boundaries between knowledge domains is the only way to achieve strong AI, and the only feasible perspective for the development of strong AI is one that integrates multiple perspectives from all relevant domains.

Thought Experiments

Albert Einstein is known for saying that "imagination is more important than knowledge." He used thought experiments as a tool to see what cannot be seen and to go where no one can go. Going into the mind of a child is not child's play. That is, building a *child* AI system is very different than building an *adult* AI system. If you want to model an adult mind and want to know what adults are thinking, you have the privilege of asking them. However, trying to figure out what a child is thinking is somewhat more complex, and more complex yet is peering into the mind of a virtual infant. The last is the crux of the thought experiment that this thesis offers.

An example of the application of this methodology would be as follows: Something in an anthropology research article sparks my attention and stimulates my recall of a psychology study, and together, these studies reinforce the results of a neuroscience experiment that I read about last week. This quickly reminds me how an engineer used those results as a guide for redesigning the cognitive architecture inside his robot. All of these elements then provide enough information to guide me through a thought experiment. In the case of this thesis, leveraging existing technology and integrating dissimilar research gives me a sense of what it is to be an infant in a virtual environment and helps me to mentally visualize the neural processes of a virtual agent engaging with its environment.

An Ecological Construct for the Virtual Autonomous Learner

Strong AI requires a construct that can build meaning in a way that resembles biological meaning building. This section articulates the key plausible elements for a virtual learning environment for the virtual student VAL. With VAL and its 3D environment, I seek to exploit embodied cognitive principles with the goal of having VAL develop strong AI status. I shall begin with an overview of my approach.

An efficient search for inspiration for the graphic user interface (GUI) and the overall design features of VAL should begin with a look at industries that have similar tools and structures. Moreover, since an embodied approach requires an environment for the body and a body for the mind, VAL's construct should be based on a multidimensional environment or 3D software platform.

The 3D software industry has been developing virtual tools and environments for over 30 years, and there is a gamut of 3D software tools for modeling, animating, rendering, and building simulations. These tools serve mammoth industries like special FX entertainment, 3D-animated films, 3D computer games, and 3D graphics for websites. From a manufacturing perspective, the 3D Computer Aided Design (CAD) industry is a robust multibillion dollar industry that provides visualization and analytical tools for engineers and architects, such as finite element analyses. These 3D tools create virtual products that are so realistic that educators use them for training pilots, soldiers, and medical personnel in life and death situations. So this thesis does not redesign the metaphorical virtual wheel; rather, it leverages different technologies with the aim of creating a new research tool that designs and develops cognitive devices.

Within the 3D-environment software landscape, the diverse community uses a multitude of interface and design conventions that provide a common foundation for anyone who wishes to build in 3D. Therefore, in order to be comprehensive, much of what is illustrated below has been drawn from these standards and conventions.

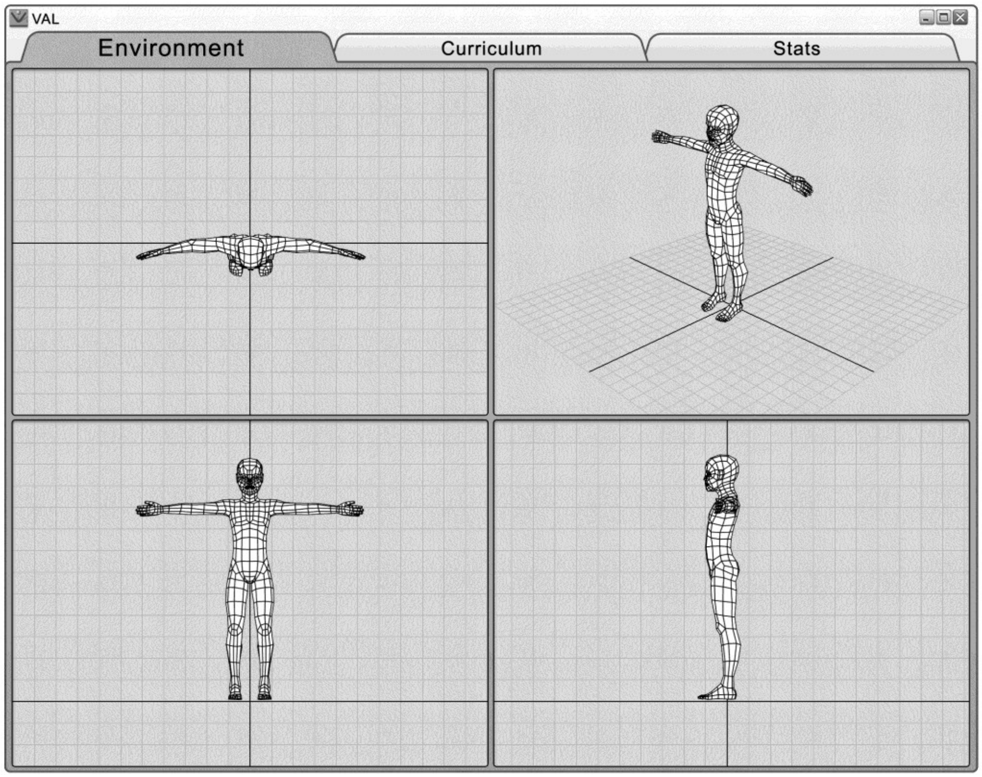

Figure 1. Sample GUI. This figure illustrates the Environment module that has a four viewport configuration. The other two modules in VAL's construct are Curriculum and Stats.

Environment

VAL's construct is, first and foremost, a 3D environment. The main interface of such an environment should have multiple viewports to help navigate within the 3D space, as illustrated in Figure 1. Standard viewing angles within these viewports consist of a perspective view, and front, back, right, left, top, and bottom views. With simple clicks on a button, a viewport can be resized or reoriented to display any view. Most importantly, a viewport can display the virtual agent's point of view. That is, the viewport can show you what the agent VAL is actually seeing at any given moment. This is a significant feature of virtual worlds, because in the real world, we never get to see a student's true perspective.

Using standard tools, researchers can build fundamentally anything within these viewports: a chair, a desk, a classroom, a school, or a humanoid infant body. The degrees of detail that these virtual artifacts have can vary dramatically, where more detail equates to a greater computational demand. The degree to which these virtual artifacts can be displayed to us, as opposed to the virtual agent, also varies dramatically. Our need to visualize the virtual content has a computational cost, and the cost is greater when we want to see more. There is a distinct difference between (1) the internal mathematical representation of objects and (2) the display resolution or what we see on the monitor. It is important to note that the viewports are used to see into and set up the environment; the agent itself does not need the viewports. The learning ramifications of this fact will be articulated in the next section, "Learning Methodology for a Virtual Autonomous Learner."

Virtual efficiency will be a recurring topic, as it is an important issue with regard to VAL's environment (or any other virtual environment). Due to the limitations of computer processor speed, there is always a compromise between realism and performance in virtual environments. For example, with today's technology, we can easily create virtual installations that are physically and photographically identical to real world installations; however, there is an efficiency cost for this level of realism, and that cost is such that most computers would not be able to navigate in real time within such a high-resolution environment. Nevertheless, significant and vital calculations can still be made based on the interaction between the environment and the agent's body, depending on how one compromises with the amount of resolution in the system. Since efficiency will be a recurring issue in this thesis, compromise will be a recurring solution.

Physics simulation. A virtual embodied approach to cognition requires the agent to interact with its environment, and this interaction should resemble real-world interactions. For example, objects have mass in the real world, and all mass is subject to a force that literally keeps us grounded: gravity. This force and other properties of the real world can be simulated in virtual environments with what is called a "dynamic physics engine."

The embodied cognitive way for VAL to learn is to learn like a real infant. Therefore, VAL's environment must have the same forces as our real-world environment. When a baby opens its hands and their contents fall down, the baby can quickly learn the effects of gravity. This is a simplified example of how a baby learns the effect of gravitational force: that all things fall down. Therefore, as long as the same forces are in place, there is the potential for VAL to learn through VE principles.

There are numerous open-source real-time physics engines that simulate the forces that are needed for a virtual agent to learn causality, that is, the relationship between an event (the cause) and a second event (the effect). Real-time physics engines create dynamic virtual environments that are similar to our real world. Virtual physics engines not only simulate gravity but also include a wide variety of other features. An important example is collision detection. Collision detection is performed by a system that defines the boundaries of virtual objects and that forces all other objects in the environment to respect these boundaries. Physics simulators also implement rigid body dynamics, soft body dynamics, and fluid dynamics. See Boeing and Bräunl (2007) for a more extensive review of open-source physics simulators.

Objects. Much of VAL's learning will come from the interaction between VAL and the virtual objects in its virtual environment. Virtual objects are just mathematical representations of real objects. Standard 3D tools can be used to create these objects within VAL's virtual environment through a process called "3D modeling." There are technicians in the 3D industry who specialize in modeling, and the created artifacts are routinely imported and exported across applications and platforms. There are also millions of premade objects that can be imported from hundreds of digital libraries. Once imported or created in VAL's construct, these objects can be modified and virtual properties can be assigned to them, giving them the characteristics of their real world counterparts.

As discussed above, the computational cost associated with the amount of details and properties to be represented is one of the main issues surrounding objects in virtual environments. The representation of a real-world object in a virtual environment is a mathematical representation that uses a virtual wire mesh to represent the volume of the object. The wire mesh can be dense or light (see Figure 2), where a less dense wire frame is less realistic but also less computationally demanding. It is important to note once again that much of the computational demand arises from the computer showing us (rendering) the results of the representation. That is, displaying the virtual environment on our monitor has a computational cost.

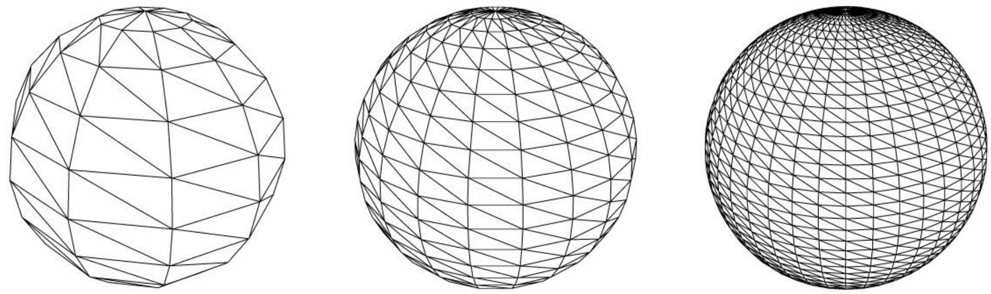

Figure 2. Different levels of wire mesh density. This figure illustrates the different levels of wire mesh density that a 3D virtual object can have: a) low-density sphere with 10 subdivisions within its vertical and horizontal axes, b) medium-density sphere with 20 subdivisions within its vertical and horizontal axes, and c) high-density sphere with 40 subdivisions within its vertical and horizontal axes.

Once the wire mesh for an object is created, it can be covered with bitmaps (image files) and texture maps, and it can be assigned other surface properties. Any image or pattern can be applied to a virtual object, making it difficult for a person to differentiate the virtual object from its real-world counterpart. Virtual objects can also be assigned optical properties like transparency and translucency that control the reflective and refractive properties of these objects.

Virtual lights, objects that behave like real lights, illuminate virtual darkness and cast shadows of objects. Through a computational process called "ray tracing," the visual effects of the lights in a virtual environment are akin to those of their real-world counterparts. Ray tracing traces virtual rays of light as they interact with the surfaces of objects in the virtual environment, and it does this in a manner that resembles real light.

The process of computer software calculating all of the variables in a scene, figuring out what the environment should look like, and then presenting it in one of the viewports so that we or an avatar can see it is called "rendering." A skilled computer 3D modeller and renderer can easily create virtual environments that are indistinguishable from real environments.

Along with optical properties, virtual objects can be assigned (tagged with) physical properties. For example, they can be soft or hard, light or dense, and fragile or durable. These physical properties work in conjunction with the dynamic physics engine. The tagged physical properties tell the physics simulator how to manipulate the objects. So, for example, a spherical object in a virtual environment that is given the properties of a bowling ball will behave like a bowling ball, and a spherical object that is given the properties of a basketball will behave like a basketball. Impressively, the rendering can be performed in real time.

Knowing that there is a computational cost to providing details, we should determine the degree of modeling details and the quantity and quality of the properties assigned to any given object based on the amount of interaction VAL would have with said objects. For example, there is no need to create an über-physically-realistic toy if that toy is on a distant shelf, far out of VAL's sight and reach. The computational cost of such an object would be high, while the learning value for VAL would be low.

Other agents and avatars. If VAL is to learn like other children, interactions between VAL and other agents are of paramount importance. The technological constructs of today's virtual environments allow such interactions, which can take on many forms. For instance, a 3D agent like Steve (Rickel & Johnson, 2000), which was discussed in the "Literature Review" section, can be preprogrammed and preanimated to guide VAL in much the same way as it would guide a human student. VAL itself is another example. When VAL is a complete agent, regardless of how profoundly its cognitive abilities have developed, it can interact with an instantiation of itself. Basically, VAL could engage in playtime with another agent like itself.

An avatar controlled by a human is probably the most robust device that can easily be implemented to facilitate VAL's learning. A virtual avatar controlled by a human can be placed in the virtual environment in several ways. For example, the avatar can have virtual volume and "physically" interact with VAL in much the same way that avatars interact with other avatars in the virtual world of SecondLife. A human-controlled virtual avatar can be coupled with forced-feedback technology. Haptic technology (tactile feedback) systems that can transmit forces between the real world and virtual worlds have been around as long as virtual worlds (Hayward, Astley, Cruz-Hernandez, Grant, & Robles-De-La-Torre, 2004). Haptic technology has great potential for physical engagement with VAL; I will discuss this further in "Learning Methodology for a Virtual Autonomous Learner."

A more efficient, but less engaging, way for a human to interact with VAL would be by appearing in the virtual environment as a hologram; however, the downside of this technique is that it would not have virtual volume. The system could be based on Microsoft's Kinect (now open-source), a virtual interface for the Xbox 360 video game console. This system uses an infrared projector and camera that can capture a human's gestures and then generate a virtual avatar in the virtual environment (Lowensohn, 2010).

Use of an off-the-shelf webcam for projecting a person into the virtual environment is another technique that is similar to the hologram and even more efficient, although less engaging. Webcam videos are routinely streamed into 3D environments, and this would be an efficient means of communicating with VAL. A video window can be placed anywhere within a virtual environment—for example, as part of a virtual monitor or on its own as a hovering virtual window.

Body

The virtual body of a 3D agent is basically a virtual robot. However, unlike real robots, a virtual robot can be created with components and techniques that are not currently available in the real world. For example, most real world robots, even humanoid robots, are constructed with mechanical parts—the parts are "hard," as opposed to organic "soft" parts. The fact that virtual environments can model organic parts is very important and cannot be overstressed. Most of this thesis is centered on embodied principles, and it is clear that the interplay between the body and the environment will vary according to the morphology of the body. Furthermore, some elements of our physical body—mostly soft elements—simply cannot be recreated in the physical real world. Obvious examples of body parts that can be created virtually, as opposed to actually, are muscle and skin.

Typically, in today's virtual environments, anything is possible. More than ever before, we have the luxury to design and develop artificial learners that conform to our research needs. To reiterate, in the real world we can only create what can technically be created, while in the virtual world, in contrast, we can create virtually whatever we need.

The virtual construction of a virtual body can be based on a mechanical model, an organic model, or both.

Since virtual bodies are virtual, they can take the form of any creature—be it a biped like us, a quadruped like a dog, a fish, or even a very simple creature like a worm. This feature of VAL's construct is important, as it allows us to study the ecology of many different creatures with different morphologies, and this provides many benefits. For example, studying virtual versions of simple worms (worm agents) minimizes the cognitive elements within the construct and therefore reduces extraneous variables. Moreover, from an evolutionary perspective, modeling a simple life form like a worm—analyzing how it interacts with its environment—may enable us to examine that creature's fundamental learning ecology. Such fundamental learning devices can then be applied to more complex agents like VAL. In fact, there are specific virtual cognitive simulators that only contain wormlike agents, for instance, the C. elegans (Kitano, Hamahashi, & Luke, 1998; Ferrée & Lockery, 1999). The important point here is that because the C. elegans has only 900 cells, of which 300 are neurons, there is little need to compromise when it is virtually reconstructed. The virtual construct for the C. elegans serves as partial proof of the concept, with VAL's construct being nothing more than a scaled-up version. I will discuss the importance of this well-researched "model organism" in the concluding section of this thesis.

Since the point of this thesis is to discuss the construction of a virtual learner that can potentially generate behavior that resembles human behavior, it seems logical to discuss the construction of a virtual robot that resembles humans. Furthermore, because this thesis calls for the development of strong AI through the development of a virtual child, VAL will be the representation of a human child.

Below, I describe what might seem like a biological system—and that is the point, because much of my design is inspired by biological systems. My idea is to try to build something that is close to a human, because we know a lot about how human parts work. Real-world *humanoid* robots, as the name suggests, resemble humans on some level. However, this resemblance is superficial. Looking beneath the surface, we see motor actuators and other components that look nothing like what we look like inside. We are forced to use these *inorganic* components because these are the only components available in the real world. Contrast this to virtual worlds where anything is possible: we can build virtual robots that look like humans inside and out, as long as we know where and how to compromise the construction of the components.

The modeling of a humanoid robot in a virtual environment can be divided into two distinct parts, although it remains one complete complex system: (1) the *Body* (discussed in this section), which is essentially the interface between the mind and the environment, and (2) the *Control System* (to be discussed in the next section), which contains the mind and controls the virtual body. This is a simple setup, for with an embodied perspective on cognition, "brains are first and foremost the control system" for bodies (Clark, 1999, p. 506).

Sensory system. Virtual senses are a key component of any complete agent because they provide the input stimulus data (perceptions). The sensory system forms an integral part of the learning mechanism and helps build meaning. The sensory system is an integrated system that includes the five well-known senses—vision, hearing, somatic sensation (touch), taste and olfaction (smell)—along with other elements that help code modalities like temperature, proprioception, pain, and pressure. In the virtual environment, different sensors will measure specific environmental values such as the intensity and color of light, the intensity and fluxuation of sound, the intensity of pressure, and the degree at which a limb is flexed. The basic function of a sensor is to gauge or perceive something in the environment and generate a value (a number) corresponding to what it perceives, outputting that value to VAL's control system.

Vision. Similar to other components of VAL's construct, its vision system is modeled on biological vision. The vision system is comprised of a virtual object (eye) that contains a grid (network) of artificial neurons that can detect light in much the same way that our photoreceptor cells detect light. Like a biological retina, the visual system can contain multiple ANN grids (cones or rods) that specialize in different optical properties. When virtual light activates the elements of the ANN grid, the grid transmits the pattern (what is seen) to VAL's control system for further processing. The actual construction of a virtual eye can vary dramatically; it can resemble the complexity of a human eye, including a functioning iris and a lens, or it can be a simplified version that only has a retina.

The importance of binocular vision with regard to cognition is well known; depth perception is one of the main cognitive benefits. The modeling of eyes provides a good example that illustrates the advantage of working within a virtual world. Once a complex bodily component like an eye (or any other complex virtual object) has been built, identical or mirror copies can be made effortlessly. In contrast, nothing about building robots in the real world is effortless. Besides the time cost that is associated with constructing complex objects in the real world, there is also a financial cost. In a virtual world, the cost of making a copy of an eye in order to give VAL binocular vision is virtually nothing.

The function of the virtual eye is relatively simple. It uses the same process that calculates and renders all of the objects in a 3D environment and then projects a 2D image of the scene on a monitor. The system calculates the position of VAL's head and eyes with respect to its environment, but instead of rendering an image on the monitor, the construct projects the same information directly into the ANN grid of VAL's virtual eyes. The data collected by the light-sensitive grid is then sent to the control system (see the section "Control System" for further elaboration).

Hearing. Hearing the sounds of one's environment is an important part of embodiment and hence an important part of learning. The physics and mathematics of acoustics form a well-established science. Acoustics engineers often use virtual simulations to map how sound will travel in real-world environments. Similar technologies can be used to generate virtual sounds from virtual objects and input the values into a virtual agent's ear. The human auditory system has two ears, one located on each side of the head, creating binaural hearing. VAL will also have binaural hearing. Even though the sound is virtual—that is, the sound is directly input to the ears—the input to each ear sensor will be different. As VAL moves within its environment, the input will constantly be updated, just as it is in the real world.

Olfaction. Olfaction, the sense of smell, is similar to hearing in the sense that it is passive; for the most part the olfactory sensor simply receives data. Similar to the way in which virtual sounds can be generated in a virtual environment, virtual odors can also be generated. Odors propagate in a much simpler fashion than sounds, and so the computational cost of such a feature is minimal, with important benefits.

In much the same way that properties are assigned to objects so that the virtual physics engine knows how to handle the objects, properties giving odor characteristics can be assigned to objects. That is, all objects can have unique numerical values that represent odor characteristics. If a tagged object is within range, the sensor located in VAL's virtual nose would pick up the information and send it to the control system.

Skeletal muscular system. The virtual skeletal muscular system contains both a structural system and an actuation system that is modeled on vertebrate anatomy; this system is therefore unlike most real-world robots. The skeleton and the muscle mass covering it make up most of VAL's morphology and therefore have a profound impact on how VAL perceives its environment. In VAL's construct, the bones and muscles of an agent, like any other object, can be of any size or shape and have any degree of detail. The purpose of a virtual skeletal muscular system, like its real-world counterpart, is to provide an agent with a support structure and movement.

Virtual bones, the hard parts of the skeletal muscular system, do have real-world counterparts—aluminum is often used as the structural frame for robots because it is light and strong. However, virtual muscles, the soft parts of the system, do not have real-world equivalents. Virtual muscles are virtual actuators and are key elements of the overall construction of VAL, mainly because they are efficient, that is, easy to construct and maintain. In contrast, electroactive polymers (artificial muscles) are still highly experimental and are not widely used. Because true artificial muscles are hard to design, most real-world robots have motor actuators.

Mouth: Vocalization and consumption. In the early stages of development, newborns and young infants have overt oral fixations. This stage of human child development is an important step as a child acquires knowledge of itself and its environment. In addition to serving as a vital passage for respiration, the mouth also has two important roles in communication and the consumption of nourishment. Not only is the mouth essential, but the upper aerodigestive tract is "the most complex neuromuscular unit in the human body" (Rogers & Arvedson, 2005, p. 1). Further evidence of this complexity can be seen in the cortical homunculus maps of the brain (Penfield, 1961), where the mouth has the second largest neural representation, after the hands. Its physical complexity, neural representation, and basic contribution to survival make the mouth one of the most important body parts. Therefore, great attention should be given to the design of a virtual mouth.

Historically, in the 3D modeling industry, the mouth (including the tongue) was notoriously difficult to render and animate, due to its complex combination of hard and soft tissue. However, this is no longer an issue; anatomically precise high-resolution faces are readily available on the market for a nominal price. In fact, there are biomechanical modeling toolkits that specialize in orofacial simulation (Stavness, Lloyd, Payan, & Fels, 2011). Such 3D models have more than enough detail for VAL.

Vocalization. Adhering to the embodiment principles, virtual vocalization should be modeled with a specific type of speech synthesis that uses the physical properties of the vocal tract. Systems using a research technique called "articulatory speech synthesis" are modeled on the human vocal tract and contain all of the variables needed to generate human utterances (Birkholz, 2010).

ArtiSynth is an open-source 3D biomechanical modeling toolkit for physical simulation of anatomical structures. In this system, both hard and soft tissue components of the vocal tract are modeled, combining both finite element method and multi-body capabilities "with an emphasis on computational efficiency" (Stavness et al., 2011, p. 2). This system is essentially an embodied approach to speech synthesis and is a testament to the advanced state of technology with regard to speech synthesis. However, even though Stavness et al. claim computational efficiency, this system can be considered a medium- to high-resolution system and therefore most likely would be computationally very costly to integrate into a complete agent like VAL.

The important point is that VAL should be given the opportunity to learn to coordinate and control the necessary variables in order to generate desired sounds. Therefore, the vocalization system need not be a high-resolution system containing all of the physical components of the vocal tract; it can be constructed as an integrated subsystem. A detailed description of such subsystems appears in the section "Learning Methodology for the Virtual Autonomous Learner." As long as the subsystem has the same variables as the high-resolution system, VAL will have an opportunity to generate sound. Therefore, a basic low-resolution virtual model is all that is needed for VAL to "find its voice."

Consumption. Consumption is a major necessity for survival and must therefore be a significant part of an embodied experience. The gut drives much of our action because the strength of our basic need for nourishment is so profound. However, a high-resolution, fully functional digestive system is not necessary. In fact, most of our internal organs are not vital to model because we are not conscious of them. Having said that, however, there is a need for a virtual agent to have a virtual representation of the gut in an embodied approach to cognition, In fact, the stomach is arguably the most relevant internal organ, and in the section "Learning Methodology for a Virtual Autonomous Learner," I will discuss how a direct connection between the gut and the control system builds meaning.

Skin. VAL's skin is an important component of the construct and contains some of the sensors that are prevalent throughout its virtual body. The various sensors allow it to touch, differentiate between hot and cold, and feel pain and pleasure.

VAL's virtual skin conforms to VAL's skeleton and muscles. The degree to which the skin conforms and stretches is determined by a process called "skinning." In the 3D animation world, there are technicians that specialize in skinning virtual characters. When constructed properly, virtual skin can behave in much the same way as human skin does. Like all of the other objects in this construct, the skin can be of any resolution. In principle, the skin covering the entire surface area of VAL could be one large somatosensory system. However, for efficiency purposes it would be prudent to limit sensors to areas that will potentially be used.

The skin is also a vital component of VAL's body insofar as the skin is what interacts with the collision-detection element of the real-time physics engine.

Virtuality. Since the canon of embodied principles with regard to cognition relies so heavily on the interplay between the body and its environment, it would stand to reason that the morphology of the body will dramatically impact its cognitive processes. This is the major argument for the development of strong AI within a virtual environment rather than a physical environment. Virtual environments contain virtual agents, and these agents are cheaper, more robust, and more diversifiable than physical agents. That is, we can create virtual agents that are just not possible in the real world.

By now you should be able to visualize all of the parts of VAL. Perhaps its virtual body is similar to one of many 3D human characters from your favorite computer-animated film. If not, try to visualize a virtual body that is complete, containing a full skeleton with eyes, ears, mouth, muscles, skin, and an array of senses integrated throughout the body. However, this picture is still missing a command and control system, which I shall cover in the next section.

Control System

VAL's control system is a sensorimotor system that acts like a brain. Here again, much of the design inspiration comes from biological systems like the mammalian brain's central nervous system. VAL's control system connects all of its virtual body parts: the sensory system, the skeletal muscular system, and the mouth. ANNs are used to make these connections. Inspired by the natural neural networks in the human brain and nervous system, ANNs consist of simple mathematical units called "nodes" and the connections between them. A single node in an ANN consists of a number of weighted inputs and usually one output. The basic function of a node is to generate an output *if* the sum of all inputs reaches a specific numeric threshold. This is similar to the firing of a biological neuron.

The nodes in an ANN are locally grouped and massively interconnected. The network connections are made by connecting the output of one node to the inputs of other nodes. Groups of ANNs can be considered to be miniprocessors that can learn from their environment through the use of sensors. When such a dynamic system is turned on, an array of sensors takes a snapshot of the environment, creating an internal state. Then, based on previous experiences from previous states, the neural network calculates the best action to take and creates the action by sending a numeric value to virtual actuators.

The basic arrangement of the ANN that comprises VAL's control system coordinates input from an array of sensors, directing the input to the dedicated processing location. Each type of stimulus modality would have its own dedicated processing location. The arrangement of the various processing locations and the path that information takes is called the "cognitive architecture" of the agent. Basically, a cognitive architecture is a metaphorical blueprint that illustrates the essential structures of and relationships between the different components that constitute the agent's thinking. There are a multitude of such architectures today. VAL's construct is set up in a manner that can accommodate a wide variety of cognitive architectures, because I believe that this approach will lead to a more productive developmental process.

In addition to coordinating the input from the sensors, the cognitive architecture generates and coordinates the output in the form of actions or behavior. Hence, the control system operates in a continuous cycle, as a dynamic system. Smith & Thelen (2003) further discuss dynamic systems. The cycle of retrieving input from sensors, processing, and then outputting to produce movement is considered to be an agent's cognitive process. The goal with VAL is to generate action that resembles human action.

Construct Summary

My basic purpose in this section was to convey the idea that virtually anything is possible in a virtual environment. Moreover, it is relatively easy to set up a virtual environment and its artifacts to look and behave like real-world environments. Modeling and configuring agent components like the skeleton, muscles, skin, and other important sensory components is also *not* an issue. Rather, the cognitive architecture of the control system is the ongoing challenge that will determine whether or not VAL will learn on its own.

So far in this thesis, what I have written about the VAL construct has been centered on the environment and its artifacts, agents, and other objects. This part of the construct is just one of three modules; the other two are a *Curriculum* module and a *Stats* module (see Figure 1). The Curriculum module is the place where learning interventions will be created and developed for VAL. These interventions will be further discussed in the next section, “Learning Methodology for a Virtual Autonomous Learner.”

The Stats module, which is not discussed extensively in this thesis, is the part of the software application that supports viewing and comparing scientific/empirical data. The quality and quantity of the data would be very acute due to the nature of virtual environments; the virtual artifacts are already digital data. As a consequence, any behavior variable, however mundane or profound, can easily be measured. For example, if VAL picks up a ball, the force exerted by each finger can be measured, along with the time it took to apply the force, the gaze of the eyes while the ball was grabbed, and so on. Given this overt amount of control over the virtual learner and its environment variables, it is easy to see how virtual learning environments have more advantages than do real-world settings. These advantages will be covered in the next section.

Learning Methodology for the Virtual Autonomous Learner

To make a small but significant point, I ask you to imagine a time in the near future when VAL is a complete agent and is similar to a virtual infant. Imagine that its construct is fully functional and that the virtual environment in which VAL resides resembles an average classroom. If your imagination takes you to a place where VAL and its environment look like a typical 3D character in a typical computer-animated film, a vision that we might call a “high-resolution environment,” then this virtual environment should contain all of the necessary educational elements that facilitate learning and thus resemble a real-world educational system. To clarify, if there are virtual schools, teachers, peers, exams, and so on *in* VAL’s high-resolution world, then it would be accurate to assume that any real-world educational lesson, activity, or intervention would and could be applied to VAL’s virtual curriculum.

I am asking you to visualize a high-resolution scene (environment and artifacts). This may, logically, lead you to ask: How important is the resolution of an embodied 3D agent? The answer is that it is very important if we want VAL to learn with ease, in much the same way a real child does; but it is not as important if we are trying to figure out the fundamental algorithms of learning within an embodied cognition paradigm. This thesis claims that the virtual resolution of VAL and its environment is a controllable variable in the development of strong AI; as long as the agent’s construction and cognitive architecture have properties and variables that are similar to those of the high-resolution version, the learning algorithms should be similar as well.

If you concur that all real-world learning paradigms would work in high resolution, then there is good reason to retain this assumption even if VAL and its environment have been scaled down to low resolution in order to meet the computational demands of the host computer. Therefore, if VAL’s learning methods include all real-world learning paradigms, it becomes more important to elucidate an *added* advantage of learning in virtual environments. The relevant question is then: How would VAL's learning differ from real-world learning?

This section has two main focal points. The first is a discussion of what can be done in virtual environments with respect to learning and development that cannot also be done in the real world. This discussion occurs in the subsections "Subsystems," "Machine Learning," and "Time Manipulation and Automation." The short section "Research, Life: Developmental Guides for VAL" contains the second focal point: it illustrates how research and other aspects of life can be used as a guide to aid the development of the virtual learner. Finally, I end with a thought experiment that merges some of these elements and in doing so illustrates further pedagogical details of VAL's curriculum.

At this point, it is important to first explicitly and precisely articulate the stage of development our virtual learner is in, with respect to its real world counterpart. I have previously stated that the development of strong AI would be more successful if it implemented Turing's (1950) idea of creating something that can simulate a child's mind. I have also alluded to my belief that a newborn with its "blank slate," which we might call the "default wiring of its neural network" or its "cognitive template," would be an ideal theoretical starting point. However, many, if not most, researchers believe that learning starts prenatally. The embodiment perspective is also in line with this belief: as long as there is a mind wired to a body that has sensors to perceive its environment and actuators to act on it, the agent has the potential to learn. As Gallagher (2005) says, from an embodied cognitive perspective the "prenatal bodily movement has already been organized along the lines of our own human shape" (p. 1). Therefore, as one of the points of this section is to illustrate how research can guide the development of VAL, it would be prudent to clearly state that VAL's cognitive template is akin to that of neonates. More precisely, my aim is to develop a cognitive template that would generate behavior that can be compared to a newborn's behavior. So, even though it is widely accepted that learning begins prenatally, it is safe to say that there are more empirical studies, and hence more data, about newborns than there are about prenatal babies. For this reason, it is just more practical, from a research point of view, to use data based on newborn development as opposed to prenatal fetal data.

Moving forward, the idea is to use the milestones of child developmental stages as guides for training and assessing VAL's ANN. Let us therefore assume that VAL's default cognitive template is striving to resemble its real world counterpart: the blank slate of a newborn. Now that I have specified that VAL's default cognitive template's age is analogous to that of a newborn, I need to say more about innateness. This is because I will soon be discussing the "programming" of subsystems within VAL's cognitive architecture, and this may at first glance seem to be the cognitive equivalent of adding innate knowledge. However, it really is not, and here is why.

It is not *that* difficult to understand the theory behind the embodied approach to cognition when observing the physical body of an animal that affords all properties necessary for survival. Here the morphology of the body *is* evidently innate; its genes determine its shape (for the most part). It is the body's interaction with the environment that drives the animal's behavior and that will ultimately determine how *well* it will survive. However, this "physical" innateness is very different from the formal philosophical doctrine of innatism that claims that ideas in the form of knowledge are innate and that somehow this knowledge is stored in our genes (Winchester, 1985). This thesis does not reject innatism, but I hope you can now see why it is not necessary for learning in the context of embodied cognition.

So, with regard to innateness, the idea is to take what we know about the overall ecology and create the subsystem as a temporary placeholder for what is known. Before going into details about how this is done, I would like to pause to articulate, from a high-level point of view, how subsystems fit within the perspective of VE. So far, I have talked about the body, the mind, and the environment without saying much about another important variable: action. I illustrate the ecological value of action with the equation

$$\text{environment} + \text{body} = \text{mind} \times \text{action},$$

in which the mind creates action that controls the body and the environment. Because the fundamental function of VAL's construct is to create all four of these variables (environment, body, mind, and action), and since we can easily create three of the four elements (environment, body, and action), this means that the only *unknown* element is the mind. With that said, let us rearrange the equation to look like this:

$$\frac{\text{environment} + \text{body}}{\text{action}} = \text{mind}$$

Yes, there is much that we do know about the mind, but it is only a small fraction compared to what we know about the environment, the body, and action. However, we can use what is known about the mind to our advantage. In fact, using what we do know to flesh out the unknown is the crux of the learning methodology for VAL. This is why we use what we know to create subsystems: subsystems are based on known behavior (actions).

Subsystems

Again, imagine VAL, this time as a complete low-resolution agent with all of the body and control system components discussed in "An Ecological Construct for the Virtual Autonomous Learner." As a complete agent, VAL has various sensors to perceive its environment and a skeletal muscular system to act on its environment. The sensors and actuators are wired to the control system. The architecture of the control system is open, that is, it can take on and implement designs from any of the numerous existing cognitive architectures being developed today. To describe the function and purpose of subsystems, I first need to articulate a simplified generic architecture within the embodiment paradigm.

The generic cognitive architecture about to be described is the wiring system that controls the low-resolution version of VAL (see Figure 3). VAL's spread-eagle posture in Figure 3 resembles that of Leonardo da Vinci's *Virtual Man.* This pose was chosen because it is the standard default position used when modelling 3D characters; it gives the modeller quick visual access to most of a character's body parts. In Figure 3, the artificial neural network, a generic architecture, is superimposed over VAL. There are three main parts to this generic architecture structure: (1) sensors, (2) actuators, and (3) the network. In this architecture, the network both processes information and is the repository of information.

Note that since my immediate purpose is to describe the function and purpose of VAL's subsystems, the exact size and type of neural network is not relevant; the subsystems can be used in conjunction with any size and type of artificial neural network.

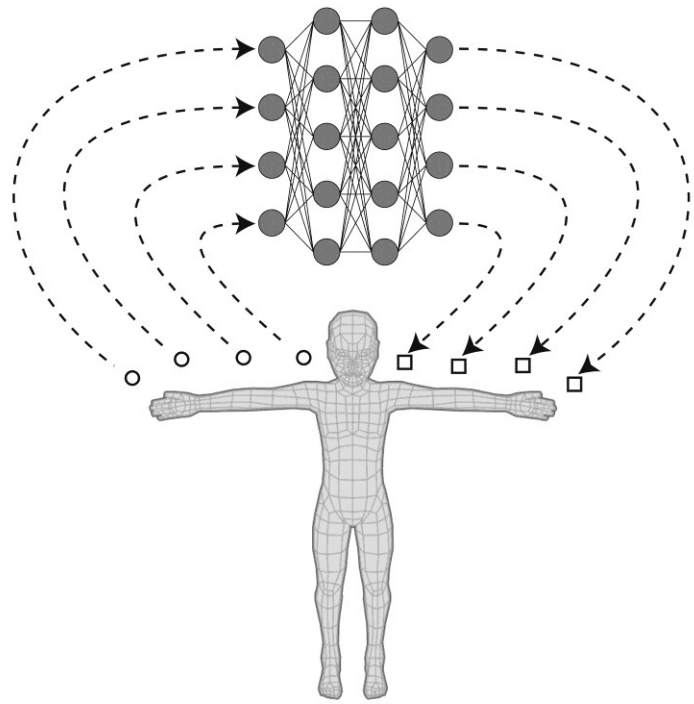

Figure 3. Generic architecture. This figure illustrates the flow of information coming from and going to VAL. Sensors are represented with circles, actuators are represented with squares, and the artificial neural network is represented by the grid.

The generic architecture is a basic dynamic input/output system with sensors and actuators creating a general feedback loop. The flow of information in this architecture is as follows. Values from all input sensors are received and aggregated within the network; let us call this network the “brain.” In the brain, the weights of neural nodes are calculated and adjusted, depending on environmental factors, and a value is output. That is, information is received from the environment via the sensors, the information is then processed as the values of the neural nodes propagate throughout the brain’s network, and finally, a value is sent to the actuators, generating action. Subsystems will, for the most part, bypass the brain's propagation process (see Figure 4). Thus, a subsystem in this thesis is an algorithmic device that is integrated into the cognitive architecture yet bypasses most brain functions—the subsystem generates actuator values based on the collective input values of the sensors. In this sense, a subsystem is a quick fix to simplify the cognitive architecture and is used as a bootstrapping device to stimulate the embodied learning process.

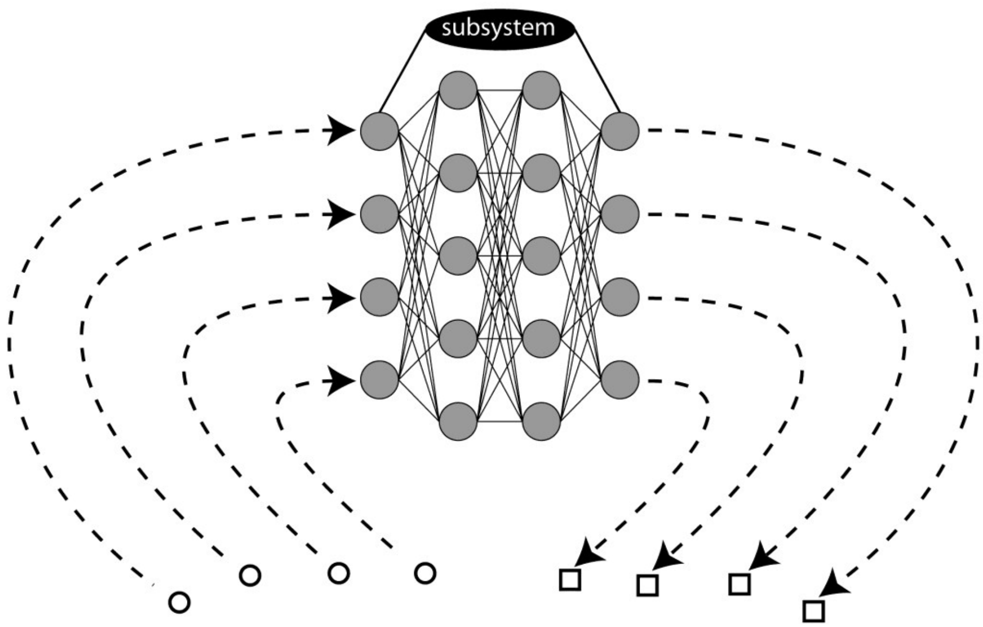

Figure 4. Subsystem. This figure illustrates how a subsystem bypasses most of the artificial neural network.

In brief, from the perspective of embodiment, the body along with the mind perceives and acts on the environment. But what initiates this engagement? From a high-level point of view, some would say that it is the basic need to survive, and this may very well be true. But what are the basic low-level devices encoded in our neural architecture that stimulate the interactions between the body and its environment? Perhaps answers to this key question can be inferred from an evolutionary investigation. Unfortunately, for the most part, this approach can only serve us at a high level. That is, evolution has shaped our bodies as well as the configurations of our neurons over millions of years, yet while we have observed and recreated the bodily shapes of past creatures, we have not done the same with their neural architectures. This does not mean that all is lost with respect to identifying and recreating the basic neural devices that are needed for an agent to learn about its environment and hence have a better chance of survival. For example, one approach would be to use behavior to infer what the neural device should "look like" and then program that behavior as a subsystem into the overall cognitive architecture, thus creating a more functional template for VAL.

To reiterate, since we know how newborns behave, or should behave, at specific stages of development, we can program very low-level behaviors that would bring about the more sophisticated high-level behaviors. The basic low-level algorithms are inferred from observed behavior and are used to design subsystems that can jump-start VAL's learning process.

For example, let us take our complete agent VAL, as seen in Figure 3. In this state, VAL is like a glorified virtual rag doll connected to a large neural network, with no initiative to do anything. However, if we program a subsystem into the motor neurons (actuators) that control eye movement to mimic human eyes' saccades (involuntary short rapid movements of the eyes), this program can act as a visual stimulus for VAL's learning because it would provide constant visual variance. The reflex actions of muscles constitute another example of what should be programmed; that is, a subsystem that mimics the reflex arcs in neural pathways should be created.

Our body's vestibular system, or our sense of balance, provides another, slightly more complex, example of where a subsystem can serve as an efficient and effective device within an agent's cognitive architecture. The bony labyrinth in our inner ears that consists of three very small juxtaposed semicircular canals is a device that helps keep our balance. Since we know how this intricate system works, that is, we know its function, there would be no point to physically and neurologically modeling this system and then recreating it in high resolution. Computationally, this would be too expensive. The efficient thing to do would be to implement a subsystem that can mimic this system's function and integrate that function into the cognitive template of VAL's artificial neural network.

The key point is that *any* system that is below the level of consciousness (that is subconscious) can be programmed with subsystems without compromising embodied learning principles. Providing such subsystems will dramatically reduce the complexity of the embodied cognitive architecture and will therefore contribute significantly to its efficiency. Our three examples of subsystems that perform the functions of the eyes' saccades, the muscular reflexes, and the vestibular system are some of the vital systems that can be preprogrammed rather than left for VAL's neural network to learn on its own. From the perspective of embodied cognitive architecture, these subsystems are just a means to control variables whose value (behavior) is known. The subsystems can then serve as a foundation on which to build more complex behaviors. This process of training and then generating the desired behaviors of computer programs is called "machine learning" and is the topic of the next section. But first, I will mention one more key subsystem.

The three aforementioned subsystems are all small devices that are nonetheless an integral part of VAL's cognitive template. However, a more complex and significant subsystem lies at the heart of VAL's learning and meaning building: a metabolism subsystem. Our metabolism consists of highly complex chemical reactions that, for the most part, convert substances into the nourishment that provides us with energy. When our energy is low and there is a strong discomfort in our gut, we have a strong desire to take action to rectify this uncomfortable state (Montague, 2006). Hence, there are two advantages to synthesizing a metabolism subsystem within an embodied agent. The first advantage is that by devising a device that can give energy to or take energy from the virtual agent, we can create an internal motivational system. The second advantage of synthesizing the metabolism is that this can create and simulate discomfort in the virtual agent's gut. In the section "Thought Experiment: A Virtual Intervention," I will explain how this subsystem is used.

Machine Learning

Now that most of VAL's inner workings have been covered, it is time to provide more details about VAL's learning methods. As I just mentioned, the training of an embodied agent with a neural network control system falls into the category of machine learning. This is a large and well-established subfield of AI that focuses on the development of learning algorithms for a wide range of computer applications (Alpaydin, 2004). Machine-learning techniques can be used with classical symbolic processing, but in this thesis I confine my discussion to neural network devices.

The core of machine learning is twofold. First, it is *descriptive*, that is, it is used to gain knowledge, where the knowledge is acquired through interaction with the environment. This knowledge is broadly distributed throughout the agent's ANN. Second, machine learning is *predictive*, that is, it is used to encourage the neural network to behave in a predictable manner. These predictions are based on experience, where more experience brings better predictions and hence more appropriate behavior (Alpaydin, 2004).

The taxonomy of machine-learning algorithms is discernible in the labels for the algorithms: "supervised learning," "unsupervised learning," "reinforcement learning," "learning association," "classification," and "regression." Note that these categories of learning algorithms are similar to those found in the vernacular of educators. Different algorithms are used in different learning situations. With this handful of algorithms, neural networks can be trained to recognize spoken words (voice recognition), recognize written words (optical character recognition), recognize faces (facial recognition), and even learn fine motor control of body parts. These learning algorithms train part of the virtual agent's neural network to respond appropriately to given stimuli. Take reinforcement learning, for example. If we were to place a virtual bottle (stimulus) in front of VAL, there is a series of events (behavior) that we would hope to happen. For instance, we would want the agent to recognize the bottle, reach for it, grasp it, pull it near, and start to suckle on it. If executed properly, the cascading chain of events would then be rewarded. The reward reinforces the appropriate behavior by strengthening the weights of the connecting artificial neurons involved in the action.

If what has been said about machine learning sounds similar to the canon of various educational learning theories, then that is a good thing, as it will help when constructing learning objectives for VAL. That is, being familiar with real-world learning theories and knowing how to teach in the real world can help researchers teach in the virtual world. However, this is where VAL's learning methods depart from real-world methods in an interesting way. The implementations of some of these learning mechanisms call for many iterations of a learning intervention to produce the desired outcome. In some cases hundreds, if not thousands, of trial-and-error cycles are needed to adjust the systemic weights of the agent's neural network to obtain the desired effect. The fact that it is so difficult to train a neural network is a good example of why these types of interventions, which we may call "nonethical" interventions, are not appropriate for humans. On a very high level, the possibility of these nonethical interventions constitutes one of the most profound advantages of virtual learning over real-world learning.

Thus, ANN learning methods are similar to human learning methods, but have some added advantages due to their ability to incorporate nonethical learning techniques. As another example, in the virtual world we can isolate any part of the agent, remove it, fiddle with it, and then replace it. This holds true for components like the agent's head, arm, skin, bones, and eyes, and even all or a portion of its neural network. The implementation of these algorithms is not only nonethical—it is impossible in the real world.

The possible of a nonethical approach to learning leads me to discuss one more type of learning algorithm that is a powerful tool for getting an agent's neural network to look and behave a particular way. Genetic learning algorithms (evolutionary algorithms) have been used with remarkable results in producing organic behavior in virtual environments (see Sims, 1994). Not surprisingly, genetic algorithms are inspired by biological evolution. These algorithms can evolve not only the physical morphology of an agent but also its neural architecture, which determines its behavior. Since it is established that the agent VAL resembles a human infant, there is no need to apply genetic techniques to its morphology. Although genetic techniques can refine the morphology of our virtual infant to breed a better baby, this is not paramount since we have more or less what we need with respect to the agent's bodily shape—that is, we know how a typical baby is supposed to look. In contrast, since we do not have a clear picture of the default configuration of newborns' neural architecture, genetic learning algorithms can be used to optimize the design of the artificial neural networks, effectively evolving a better cognitive architecture and hence better default cognitive templates for VAL.

To reiterate, the idea is to evolve the neural architecture and the weights of its nodes in order to make the architecture more likely to generate appropriate behavior. Since VAL's behavior is at issue, the genetic algorithm discussed here applies only to the agent's artificial neural architecture and not the shape of its body. A genetic algorithm in machine learning is an evolutionary process that filters out bad designs, specifically, bad neural designs. The algorithm finds design solutions for a specific task by using a fitness function to select specific properties of the neural network that make it more inclined to perform the task. This is a Darwinian approach that generates multiple potential solutions in the form of multiple neural architectures and/or neural weight values and that then applies a fitness function to select the best candidates. The top contenders (neural designs) are then combined and used to generate new candidates. The process is reiterated until the neural architecture produces the desired results, which in this context are appropriate behaviors or appropriate actions with respect to the environment. Variations of this technique have had remarkable success in developing behavior in virtual environments. For example, complex control systems have been designed (evolved) to enable agents to walk and jump (Streeter, 2003).

I have said that the use of nonethical learning methods is one of the most profound advantages of virtual learning over real-world learning. Now it is time to introduce the most "technical" advantages that the virtual world has to offer: time manipulation and automation. These topics are discussed in a separate section since they have more to do with the nature of computing then the learning process itself.

Time Manipulation and Automation

Time in the real world is limited for students and their education. There are only 24 hours in a day and only 365 days in a year. Perhaps educational policymakers take this limited duration into consideration when designing and developing curricula. In fact, time is such an integral factor in learning and development that there are time-based markers that gauge whether or not a child's "mind" is emerging at a normal rate. Furthermore, it is probably safe to say that most educators think that "time on task" is an important variable in a student's acquisition of knowledge. Since it is clear that time is a key element in development and learning, *virtual time* is a valuable asset for the curriculum and pedagogy of a virtual agent.

Like most other aspects of the virtual world, virtual time can be manipulated. Most notably, virtual time (1) can be compressed and (2) can be used to manipulate the energy level of agents. With regard to time compression, imagine a three-hour virtual lesson being compressed into five minutes of real time, or years of virtual learning being compressed into days of real time. All of this is possible. I will develop this idea further in just a moment, but first, I shall say more about the manipulation of the agent's notion of time.

For the most part, time for the virtual agent depends on the processing power of its host computer. VAL's construct that is described in this thesis is just a computer program that runs on top of an operating system; the program's commands are converted into 1s and 0s so that the computer's central processing unit (CPU) can process this information. The very nature of the hardware/software relationship is that commands are executed as fast as is technically possible. Programs will process information at different speeds that depend on the power of the CPU. Different computers equate to different speeds .

An embodied agent cannot have such a chaotic ecological environment, where on one system it is moving faster and on another, slower. Hence, the notion of virtual time within the agent's environment is controlled by software. In this case, the dynamic physical engine acts like a governor and grounds time to simulate real-world physics. For example, since we know precisely how long it should take for a dropped ball to hit the floor after being released from a specific height, that value can be and is used to calibrate virtual time, making virtual time a constant across all computer platforms.

With that said, it should be clear that virtual time is an added advantage of virtual environments, as it can be employed to serve learning and developmental needs. As I have often mentioned in this thesis, designing an efficient agent is paramount, and here is another reason why: in building a low-resolution agent that can "freely" interact with its environment in real time, we have room to manipulate its energy level. This energy variable is one of the key elements that is integrated into VAL's metabolism subsystem.

From a behavioral point of view, our model can control whether the agent is sluggish or energetic. Furthermore, the ability to control an agent's energy level can lead to the development of more complex feelings such as basic emotions (Greenspan & Shanker, 2004; Montague, 2006).

With regard to time compression, one pedagogical approach would be to create interventions that target specific learning objectives and then render the activities (the learning process) without displaying them in a viewport. If set up properly, all of the variables needed for the agent's behavior to be modified as a result of the interventions would still be intact, thanks to the environment's physical engine, and thus the agent's perspective on time would be effectively compressed relative to our time.

Developing the comparison to educational technology further, we can say that creation of these interventions is akin to the creation of learning objects. Moreover, it is not difficult to imagine a system where the curriculum and the applied pedagogy are totally automated. When this becomes reality, the early stages of infant development and eventually the acquisition of higher cognitive functions can be automated, effectively automating an agent's entire education. Setting up such a system is a key element of VAL's construct and is part of the rapid-prototyping process mentioned in the introduction to this thesis.

The main function of a rapid-prototyping process is to quickly create something—anything—where it is paramount to have something tangible to work on; that is, we first create and then improve. The ultimate goal in this case is to automate a series of learning interventions (dozens, if not hundreds) with the intention of having the virtual agent reach an infant milestone marker.

Research, Life: Developmental Guides for VAL

Standard AI techniques based on standard cognitive science have a limited potential for meaningful comparison with human research studies because the standard techniques lack an embodied perspective (Shapiro, 2011). Even standard connectionist models fall short in this respect. It is true that the traditional connectionist approach has been useful for explaining and understanding a wide range of human cognitive phenomena (Chen & Verguts, 2010; Shultz, 2003; Sun, 2007); it is a common practice to compare human phenomena to computer simulations in order to better understand the human condition. However, taking the computational methodology one step further, embodied approaches have been able to supplement current empirical data and to answer research questions that the traditional connectionist approach cannot answer (Rucinski, Cangelosi, & Belpaeme, 2011). This should come as no surprise, as the embodied approach is more encompassing, accounting for more of the variables that relate and encode multimodal learning.

As mentioned in the section "Educational Rationale and Context," the ultimate goal of my research is to create an AI educator through the development of a virtual child. The purpose of this section is to articulate how research on child development and other aspects of human life can facilitate the development of VAL. The example I use to illustrate the research point is an experiment called "sticky mittens," and the example I use to illustrate how aspects of human life can be used to facilitate VAL's development centers on physical well-being.

With regard to child development, many researchers believe that there are cognitive windows of opportunity in which emerging cognitive devices need to be cultivated in order to mature properly. These critical periods can be viewed as embodiment issues, as they (1) are "especially sensitive to input of the environment" (Needham, Barrett, & Peterman, 2002, p. 279) and (2) are associated with many input sensory devices. For example, if an infant is deprived of proper visual stimulation within its critical period, the "visual cortical neurons" become "unresponsive to subsequent visual stimulation" (Rittenhouse, Shouval, Paradiso, & Bear, 1999, p. 347). Gathering data concerning the occurrence and duration of these windows is the first step in learning their effects on development. Thus, empirical data are documented and aggregated to determine the milestones of child development, and these milestones are regularly used as empirical points of reference.

The milestone connected with the sticky mitten research is set at "approximately 5 months of age," when "infants systematically reach for objects" (Needham et al., 2002, p. 281). In this study, soft fleece mittens covered with Velcro tabs (hard side) were used to stimulate infants' prehension skills. In brief, two- to five-month-old infants were given "enrichment sessions" ten minutes a day for two weeks. During these sessions, the infants were attired with their sticky mittens and placed in front of a table that contained attractive toys laced with Velcro strips (soft side). The researchers wanted to see "what effect the enrichment of infants' typical early experience as agents acting on objects would have on their object exploration behavior" (Needham et al., 2002, p. 280).

The researchers had a variety of trials that measured (1) visual exploration, (2) swats preceded by visual contact, (3) exploration percentage, and (4) switching between looking and mouthing. The result was that "on almost every measure of object exploration and object-directed action obtained, infants who had the 2-week enrichment experience with the sticky mittens significantly outperformed their counterparts who did not have this experience" (Needham et al., 2002, p. 290). Intuitively, these results should come as no surprise, as any attention given to a child will inevitably stimulate learning. The relevant point here is that the general design of this study can be implemented within VAL's construct, and since the general research designs of some human studies can be applied in virtual environments, the data generated can more easily be compared to human data.

As previously mentioned, *time* is somewhat arbitrary in virtual environments, and age is therefore also arbitrary. In order to gauge VAL's developmental progress, we need to have some point of reference. We need a standardized system with which to gauge performance. A system that measures comparative age would be useful. For example, VAL 0.1 would have 10% of the cognitive function of a human neonate and VAL 1.0 would be the equivalent of a newborn, and the agents' performance can be measured accordingly.

With an embodied approach to AI, it is not sufficient to pass the Turing test. That is, mastery of a test helps very little when the goal is to learn from the journey. We need milestones to gauge the learner's progress, so that we can learn from it. If the virtual artificial learner's results are similar to human trials, this will serve to indicate that we are heading in the right direction with regard to VAL's development.

Research and theories about child development are the logical places to find data that can help guide VAL's learning. A VE approach to physical development offers more relevant tools. That is, technologies that are related to physical therapy can offer techniques and tools to aid VAL's physical learning. For instance, a set of parallel bars, a commonly-used physical therapy device that supports people while they gain strength and balance, can also be used in a virtual environment. Virtual rigs can also be set up to function as leg braces; as the constraints of the braces limit the movement of the agent's virtual legs, the ANN's weights associated with the legs will also be constrained. In fact, any real-world physical device that aids people or animals can be used virtually.

As VAL's comparative age increases, more and more real-world activities can be integrated into its curriculum. For example, when there is a VAL 2.0, sports and dance can be used as tools for learning better coordination, and when there is a VAL 5.0, learning yoga would perhaps make it more aware of its body.

Thought Experiment: A Virtual Intervention

The goal of this thought experiment is to guide you through a typical learning intervention from an embodied cognition perspective and then to articulate what VAL is learning through its neural network as it interacts with its environment. At this point, it is important for the reader to consider and envision VAL as a complete agent, with all modules of its construct fully functional and doing what they were designed to do. For instance, the Stats module containing the databases can present the measureable variables in a meaningful way, the Intervention module can create any intervention you can imagine, and the Environment module can host every virtual educational need.

This thought experiment is something like a mock case. In this scenario, you (the reader) will act as a researcher going through various steps of creating and implementing an intervention and then measuring its effectiveness by comparing its results with that of a control. With this thought experiment, we revisit the sticky mittens study. However, this particular implementation can only be executed virtually. Needham et al. (2002)

tested the effects of the sticky mittens by measuring how much, if at all, they would enhance an infant's ability to grasp objects. Therefore, in this thought experiment, one of the variables being measured will be VAL's set of prehension skills. The similarities to Needham et al. end there, as the toys are replaced with a baby's bottle containing virtual nourishment. At the risk of sounding overly dramatic, I will stipulate that if VAL does not learn how to grasp the bottle, it will die. This dramatic scenario is intended, in part, to emphasize the point that anything is possible in a virtual environment, including nonethical intervention.

In the center of the vast virtual environment lies a crib containing the virtual autonomous learner. Double clicking on the magnifying glass icon causes the crib to zoom in. VAL is in a spread-eagle position, still and lifeless, since the simulation is not running—that is, VAL is turned off. This version of VAL, which we shall call "version 0.1," has an unweighted default cognitive template because this is its first intervention, perhaps the first of many in a long and complex curriculum. For all intents and purposes, VAL 0.1 is like a newborn with no life experience. It has no knowledge about its environment or even about itself, for it has not yet interacted with its environment. Therefore, running the simulation (turning on VAL) at this point would result in the agent just flailing about, owing to the fact that none of the neural weights in its artificial neural network have been adjusted (fine-tuned).

A click on the "Intervention" tab brings up the Intervention module. Here you set up the parameters for the experiment, in this case inserting a standard "rig" to hold the baby's bottle. This module is designed to keep track of all variables in the environment, but more importantly, it allows you to select and present only the variables that you want to measure: variables that measure VAL, all artifacts, all environmental properties (light and sound), and all aspects of the aforementioned rig.

Since our goal in this experiment is to measure whether the sticky mittens are an effective tool to use for refining VAL's prehension skills, we need two sets of trials: one with the sticky mittens and one without. In each trial, VAL will have to learn to reach, grab, and then bring the bottle to its mouth. VAL has a limited time to do this, since all bodily movement expends energy and if VAL runs out of energy, it will die. Survival thus requires VAL to obtain energy by bringing the bottle to its mouth, and the agent will survive if it masters this task. It is important to add that slightly touching the bottle will give VAL some energy, perhaps just enough for VAL to try a few more times. The reality of the situation is that it will take hundreds of iterations for VAL to master the task, where each new iteration follows the death of a VAL that was unable to master it (this is a nonethical approach). Given this design, one measurable variable will be the number of iterations. Supplementing this data, there will be a set of heuristics based on the quality of VAL's behavior, for example, the path that the hand took to reach the bottle, the strength of the grip, and the position of other body parts.

After all of the parameters for the intervention are set up, we click on the "Environment" tab to get back to the Environment module. Now we see the intervention rig attached to the crib and hanging over VAL, within arm's reach. For the first trial, let us start with the intervention without the use of the sticky mittens. To speed up the experiment, we would normally run it in the background where it can compress virtual time, but first let us see what happens in real time.

Clicking on the “Run in Real Time” button, we see the virtual infant flailing about, as expected. Since its neural network has not had any training, all limb and joint actuators have been given random values as weights, which are meaningless. Note, however, that even though VAL’s movements are jerky, the behavior is in real time and seems to obey the fundamental laws of physics.

Looking at the viewport’s heads-up display, we can see that VAL’s energy level is low. If VAL does not find the bottle and obtain more energy it will die, and since we want to see whether the automation process is functioning properly, we let it exhaust its energy. At this point, VAL stops moving. After the cycle is complete (death), the construct sends relevant data from this first iteration to a database.

Clicking on the “Stats” tab for a quick peek to see if the data is in order, we discover just that. Upon returning to the Environment module, we see that the construct’s automation process has loaded the second iteration, a second cycle in which a new VAL replaces the old. The morphology of VAL remains the same for each cycle. However, the computer randomly changes the weight values in its neural network, and these slight changes in the network produce slightly different behavior.

Let us say that the first cycle took five minutes, an arbitrary amount of time. If it takes three hundred cycles for VAL to master the task, then it should take over a day to execute this first trial. So, since we see that all aspects of the construct are functioning properly, we click on the “Run in the Background” button. The virtual environment goes black and an indicator pops up, stating that the virtual time is being compressed by ten to one. Now, if it takes a hundred cycles, all we have to wait is a little under an hour.

One hour and 286 cycles later, VAL has mastered the task. Another quick peek at the Stats module reveals that the database contains 286 entries. Looking closer at the last entry, which contains the data on VAL's mastery of the task, we see that it took just over one minute for VAL to grasp and suckle the bottle. If we now used this version of VAL with the "trained" neural network as the default cognitive template, we would find that more often than not, VAL 0.1.286 (as we call this version) can complete the task in just over a minute. The reason is that somewhere within its neural network there are weights (of neural connections) that result in a greater aptness for grabbing the bottle.

The same process used in the first trial (VAL without sticky mittens) is used for the second trial (VAL with the sticky mittens). Not surprisingly, it takes only 204 cycles for VAL to master the task in this case. Let us take a closer look at the details of the intervention from an embodiment point of view, to understand what is happening to VAL's neural network when it interacts with its environment.

Running the construct in real time and zooming in even further, we now see the familiar ungraceful movement of the virtual infant. Then, simply by chance, and at the same time inevitably, the infant's hand touches the bottle. When this happens, the infant receives a slight reward in the form of more energy. The associative learning algorithm within VAL's neural network then evaluates that bodily state (the position of the body, especially the agent's gaze and the position of its hand touching the bottle) as being a positive thing, essentially reinforcing the behavior.

The reinforcement of VAL's actions acts like a constraint, as it increases the likelihood of that action being repeated. VAL's behavior becomes more refined each time its hand touches the bottle. Each time the hand touches the bottle, it does so in a slightly different state (position), and because of this, the neural representation gets a better "feeling" for where the bottle is. The chances that VAL's *palm* will hit the bottle in the future are greater. Taking this thought experiment one step further, let us imagine what happens when VAL's palm *does* touch the bottle.

At this point, the built-in subsystem in the neural network kicks in and causes all actuators that control the fingers to clench, effectively gaining a firm grip on the bottle. When this happens, a "body state" that is associated with all the variables used in the grabbing action is created within the network. While VAL is in this state, very interesting things begin to happen, from the embodiment point of view. For the first time, VAL is really engaged with its environment. The physics engine is simulating a slight resistance that provides valuable feedback for VAL and that thereby adds another layer of modality to the learning intervention. This added modality is the sensation of "pressure" that is virtually created through the sensors lining VAL's fingers and palm. For VAL, the sensations are real because the physics engine simulates the pressure. Some other modalities contributing to the body state are VAL's gaze and its proprioception.

At this point, it should be clear that the embodied process builds on top of VAL's prior success. The sticky-mittens intervention would be just the first of many interventions lined up in a queue that make up VAL's virtual curriculum. Furthermore, the curriculum could be peppered with human/VAL interactions in which the rig that holds the bottle can be replaced with one of many force feedback devices. Then you could actually feel the pull of the virtual infant's grip.

Segments of VAL's curriculum could be automated, creating virtual milestones. Perhaps, when many milestones have been reached, that version of VAL could be saved and tagged as being the equivalent of a one-month-old baby. VAL with a cognitive template of 0.2, for example, would have the equivalent of twenty percent of an average human infant's cognitive development. The goal for all of the milestones would be to increase this percentage, thus making VAL's behavior that much closer to human behavior, and this in turn would enable us to set up interventions that more closely match real-world experiments.

Much of our discussion has illustrated VAL at an early stage of development. Because most of the virtual agent's acquired skills at this stage of development are directed at gaining control of its virtual body, let us call this a "sensorimotor stage," drawing a parallel to Piaget's (1977) theory of cognitive development. At this stage, much of VAL's development can be automated. The intervention envisioned in the thought experiment would be one of hundreds of interventions needed for VAL to advance to the next stage of development, whatever that might be. In fact, there are numerous theories of child development that can be used as a reference or for inspiration.

Discussion

In this discussion, I will first try to reinforce the premise of this thesis by asking and then answering questions related to the viability and purpose of my research. I shall then come full circle by articulating how an AI educator would be integrated into our future culture, concluding with a few final words regarding Moore's Law and AI.

All that has been discussed is just the beginning of VAL. With VAL 0.1, I introduced you to the ecological construct where VAL was created. Then, based on the embodiment perspective, I discussed some of its learning methodology. Although there were many key elements, I have boiled them down to three.

First, the construct for VAL needs to accommodate different cognitive architectures if we are to make full use of its methodology; much excellent research is being carried out in the realm of embodiment, and the ability to implement different algorithms will definitely be an asset that can contribute greatly to the development of VAL. Second, another vital element of this thesis is the development of a rigorous virtual curriculum and the design and development of efficient pedagogical tools to implement that curriculum. Third, this thesis emphasizes the importance of having an educational perspective. These are all key elements for the development of strong AI. But how much of this is possible?

To answer the question "Can intelligent behavior emerge from within VAL's ecological construct?" I refer the reader back to the "Literature Review" section. If we look at other research such as that presented in Metta (2010), Cotterill (2003), Brooks (1991), Pfieifer (2007), and even Goertzel (2007), we can see that many elements similar to those involved in the design of VAL are already being implemented. Moreover, if we view the C. elegans simulations in Kitano and Luke (1998) as being "ultra-low resolution" versions of VAL, then the answer to the question is most likely "yes."

Perhaps, because the aforementioned research is similar to mine, a more relevant question should be asked: How does my research differ? The answer to this question lies in the desired end product of the research. In the case of the CyberChild (Cotterill 2003), the goal is consciousness, and in the case of Goertzel (2007), the goal is artificial general intelligence (AGI). Moreover, although the theoretical foundations of Brooks (1991), Pfieifer and Grand (2007), and Metta (2010) are similar to mine, I am overwhelmingly in favor of a virtual approach. This has many implications.

For example, as stated earlier, there are no limits to what can be virtually built. Since the real world is more constrained, there are investigative techniques that only can be executed virtually. The main feature of my research that sets it apart from other research is the method of virtually combining the child approach to strong AI with an evolutionary approach. This method can be summarized as follows.

Start by creating a simple virtual creature, one that resembles one of our evolutionary stages, perhaps something resembling the C. elegans. Then acquire data on its cognitive architecture by creating interventions based on its interaction with its environment. Use that data to build more sophisticated model animals from our evolutionary history, for example, a fish, then an anthropoid, then a quadruped, and so on, until we build a biped. At every stage, the virtual creature (agent) learns how to survive on its own. The learning is based on embodied cognition principles, the creatures learning the affordance of themselves and their environment.

With this approach we can build virtual agents like VAL 0.1, which was used as a point of reference in my thought experiment. But how do we get from VAL 0.1 to an AI educator? The first step is getting VAL *0.1* to VAL *1.0*, which is a biped that can survive on its own. The next step and all future steps integrate a more and more sophisticated virtual curriculum to bring out more and more humanlike behavior. It is vital that the agent's behavior resemble human behavior, because students are human and live in a world that is designed (physically and mentally) for humans.

From a physical perspective, as we look at all the artifacts that surround us, it is apparent that our world is built for and by bipeds with very dextrous hands. With regard to our mental faculties, there are social structures that can only be negotiated through human capacities. It is clear, then, that an agent developed with human behavior would make a better educator than one that was not so developed, for it is important that educator and students be able to relate to one another.

I see the future AI educator as one that can talk to students without prejudice and that is available on demand. The AI educator would also function as an assistant and inevitably be integrated into mobile technology, in much the same way as Apple's Siri. In fact, I predict that in the near future (within the next ten years), every mobile device will have access to strong AI. At first, the strong AI agent will look and act like Siri, but then it will mature to act more like a real personal assistant or companion. When that happens, the strong AI agent will have become an AI educator. As it becomes more and more knowledgeable and gains more and more experiences, the AI educator could perhaps become an AI mentor. I dare to take this flow of thought one step further: I claim that one day, virtual companions will even walk among us.

Soon material technology will offer products that can be used as artificial muscles. In combination with 3D printer technology, this will enable us to manufacture robots that are just not possible today. These new robots will be manufactured with a morphology that can be identical to any virtual agent, giving us the ability to transfer the AI educator from the virtual world into the physical world.

This may seem far-fetched, but the synergy of new technology has always been an accelerant. New technology needs new benchmarks. Since I proclaim that VE is the new AI, perhaps we need a new Turing test, as suggested by Mueller and Minnery (2008). Keeping in line with human research, a simple VE test could be the "rouge test," otherwise known as the "mirror test" (Keller et al., 2004). This simple test is used to evaluate whether a child or animal can recognize its own reflection in a mirror, and it could be applied to both VE agents and physical companions. An agent that successfully passes this test could be considered conscious, which brings the ethical question “What is life?” to the forefront.

I have no intention of going into details about the ethics of robots, but an interesting caveat is in order. I would not be surprised if one day some research ethics committees were to see VE agents roaming around in their virtual environments and think to themselves that it might be a good idea to include VE research in their ethical reviews. Why would reviewing the ethics of VE be a good idea? It would be like an insurance policy covering us humans, just in case VE agents do one day become truly autonomous. If that happens, how would we look if we totally disregarded ethics when creating them? A better question, perhaps, might be: What would be the ramifications of creating such creatures?

My final words about Moore’s Law involve a comparison of processing demands as we move from AI to VE. I claim that increased computational power does little to improve the results of the old AI paradigm. To some this might seem obvious, but for others I offer a simple example, hoping to show that the more computationally demanding a task is, the more it will benefit from Moore’s Law.

This is something of an exaggeration, but I want my point to be clear. Thirty years ago, when calculating 2×2 on a new state-of-the-art calculator, one would receive the answer 4 in less than a second. Today one can still receive the answer 4 in less than a second. In this situation, because the calculation is not very complex, no real benefit arises from using faster processors. That is, not much was gained from Moore’s Law. In contrast, a more computationally demanding task like sequencing the human genome can benefit dramatically. Decoding the first billon base pairs took three years, but with today's processors, a person's DNA can be sequenced in a couple of weeks. The benefits of increased processing power for the simple calculator are negligible compared to the tremendous benefits for the more complex DNA sequencing. Similarly, since VE is more computationally demanding then AI, it will surely benefit more.

The two main points I wish the reader to take away from this thesis are: Virtual environments can serve as a productive test bed for embodied cognition research, and within this paradigm, many of the issues with regard to the development of a virtual artificial learner are educational.

References

Allen, P., & Greaves, M. (2011). *The singularity isn't near*. Message posted to http://www.technologyreview.com/blog/guest/27206/

Alpaydin, E. (2004). *Introduction to machine learning*. Cambridge, MA: MIT Press.

Beck, J., Stern, M., & Haugsjaa, E. (1996). Applications of AI in education. *Crossroads*, *3*(1), 11–15.

Best, B. J., & Lebiere, C. (2006). Cognitive agents interacting in real and virtual worlds. In R. Sun (Ed.), *Cognition and multi-agent interaction: From cognitive modeling to social simulation* (pp. 186–218). New York, NY: Cambridge University Press.

Birkholz, P. (2010). *About articulatory speech synthesis*. VocalTractLab. Retrieved from http://www.vocaltractlab.de/index.php?page=background-articulatory-synthesis

Boeing, A., & Bräunl, T. (2007). Evaluation of real-time physics simulation systems. *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia* (pp. 281–288). Perth, Australia.

Boulos, M. N. K., Hetherington, L., & Wheeler, S. (2007). Second life: An overview of the potential of 3-D virtual worlds in medical and health education. *Health Information & Libraries Journal*, *24*(4), 233–245.

Brooks, R. A. (1991). Intelligence without reason. *Proceedings of the International Joint Conference on Artificial Intelligence* (pp. 569–597). Sydney, Australia.

Brooks, R. A., Breazeal, C., Marjanovic, M., Scassellati, B., & Williamson, M. M. (1999). The Cog project: Building a humanoid robot. In C. L. Nehaniv (Ed.), *Computation for metaphors, analogy, and agents* (pp. 52–87). New York, NY: Springer-Verlag.

Chen, Q., & Verguts, T. (2010). Beyond the mental number line: A neural network model of number–space interactions. *Cognitive Psychology*, *60*(3), 218–240.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, *3*(9), 345–351.

Cotterill, R. (2003). CyberChild: A simulation test-bed for consciousness studies. *Journal of Consciousness Studies*, *10*(4-5), 31–45.

Gamez, D. (2008). Progress in machine consciousness. *Consciousness and Cognition*, *17*(3), 887–910.

Edelman, G. M., & Mountcastle, V. B. (1978). *The mindful brain: Cortical organization and the group–selective theory of higher brain function*. Cambridge, MA: MIT Press.

Ferrée, T. C., & Lockery, S. R. (1999). Computational rules for chemotaxis in the nematode C. elegans. *Journal of Computational Neuroscience*, *6*(3), 263–277.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford, England: Clarendon.

Gamez, D. (2008). Progress in machine consciousness. *Consciousness and Cognition*, *17*(3), 887–910.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gladwell, M. (2005). *Blink: The power of thinking without thinking.* New York, NY: Little Brown.

Goertzel, B., & Pennachin, C. (Eds.). (2007). *Artificial general intelligence*. New York, NY: Springer.

Good, T. L. (1981). Teacher expectations and student perceptions: A decade of research. *Educational Leadership*, *38*(5), 415–422.

Google (2012, February 1). *Company*. Retrieved from http://www.google.com/about/company/

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love & the meaning of life*. London, England: Bodley Head.

Gopnik, A. (2011). *What do babies think?* TEDGlobal 2011, Filmed July 2011, Posted October 2011, retrieved from http://www.ted.com/talks/alison_gopnik_what_do_babies_think.html

Greenspan, S. I., & Shanker, S. (2004). *The first idea: How symbols, language, and intelligence evolved from our primate ancestors to modern humans*. Cambridge, MA: Da Capo Press.

Hawkins, J., & Blakeslee, S. (Eds.). (2004). *On intelligence.* New York, NY: Times Books.

Hayward, V., Astley, O., Cruz-Hernandez, M., Grant, D., & Robles-De-La-Torre, G. (2004). Haptic interfaces and devices. *Sensor Review*, *24*(1), 16–29.

Jaynes, J. (1990). *The origin of consciousness in the breakdown of the bicameral mind*. Boston, MA: Houghton Mifflin.

Keller, H., Yovsi, R., Borke, J., Kartner, J., Jensen, H., & Papaligoura, Z. (2004). Developmental consequences of early parenting experiences: Self-recognition and self-regulation in three cultural communities. *Child Development*, *75*(6), 1745–1760.

Kitano, H., Hamahashi, S., & Luke, S. (1998). The perfect C. elegans project: An initial report. *Artificial Life*, *4*(2), 141–156.

Kurzweil, R. (1999). *The age of spiritual machines: When computers exceed human intelligence.* New York: Viking.

Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York, NY: Penguin.

Lockery, S. R. (2011). The computational worm: Spatial orientation and its neuronal basis in C. elegans. *Current Opinion in Neurobiology*, *21*(5), 782–790.

Lorenz, B., & Barnard, E. (2007). A brief overview of artificial intelligence focusing on computational models of emotions. *2007 5th IEEE International Conference on Industrial Informatics* (pp. 76–123). Vienna, Austria: IEEE.

Lowensohn, J. (2010, November 30). *Kinect's open-source ambitions*. CNet: News, 149. Retrieved from http://news.cnet.com/8301-10805_3-20023614-75.html

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W. H. Freeman.

Metta, G. (2010). The iCub humanoid robot: An open-systems platform for research in cognitive development. *Neural Networks*, *23*(8-9), 1125–1134.

McCarthy, J. (2008). The well-designed child. *Artificial Intelligence*, *172*(18), 2003–2014.

Montague, R. (2006). *Why choose this book? How we make decisions.* New York, NY: Dutton.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain, 120*(4), 701–722.

Mueller, S. T., & Minnery, B. S. (2008). Adapting the Turing test for embodied neurocognitive evaluation of biologically-inspired cognitive agents. *Proceedings from the 2008 AAAI Fall Symposium* on *Biologically Inspired Cognitive Architectures* (pp. 117–126).

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills: Early simulated experiences reaching for objects using 'sticky mittens' enhances young infants' object exploration skills. *Infant Behavior and Development*, *25*(3), 279–295.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

Penfield, W. (1961). Activation of the record of human experience. *Annals of the Royal College of Surgeons of England*, *29*(2), 77–84.

Pfeifer, R., Bongard, J., & Grand, S. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, MA: MIT Press.

Piaget, J. (1977). In H. E. Gruker & J. Vonèche (Eds.), *The Essential Piaget*. London, England: Routledge & Kegan Paul.

Rickel, J., & Johnson, W. L. (2000). Task-oriented collaboration with embodied agents in virtual worlds. In J. Cassell, J. Sullivan, S. Prevost, & E. F. Churchill (Eds.), *Embodied conversational agents* (pp. 95–122). Cambridge, MA: MIT Press.

Ridley, M. (2003). *Nature via nurture: Genes, experience, and what makes us human*. New York, NY: HarperCollins.

Rittenhouse, C. D., Shouval, H. Z., Paradiso, M. A., & Bear, M. F. (1999). Monocular deprivation induces homosynaptic long-term depression in visual cortex. *Nature*, *397*(6717), 347–350.

Rogers, B., & Arvedson, J. (2005). Assessment of infant oral sensorimotor and swallowing function. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(1), 74–82.

Rucinski, M., Cangelosi, A., & Belpaeme, T. (2011). An embodied developmental robotic model of interactions between numbers and space. *Annual Conference of the Cognitive Science Society* (pp. 237–242). Austin, TX: Cognitive Science Society.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424.

Shapiro, L. (2011). *Embodied cognition*. New York, NY: Routledge.

Shultz, T. R. (2003). *Computational developmental psychology*. Cambridge, MA: MIT Press.

Sims, K., (1994). Evolving virtual creatures. *Computer Graphics* (Siggraph '94 Proceedings), 15–22.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences*, *7*(8), 343–348.

Stavness, I., Lloyd, J. E., Payan, Y., & Fels, S. (2011). Coupled hard-soft tissue simulation with contact and constraints applied to jaw-tongue-hyoid dynamics. *International Journal for Numerical Methods in Biomedical Engineering*, *27*(3), 367–390.

Streeter, T. (2003). *Autonomous virtual humans*. Tyler Steeter website. Retrieved from http://www.tylerstreeter.net/#AutonomousVirtualHumans

Sun, R. (2007). The importance of cognitive architectures: An analysis based on CLARION. *Journal of Experimental & Theoretical Artificial Intelligence*, *19*(2), 159–193. doi:10.1080/09528130701191560

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, *9*(4), 625–636.

Winchester, S. J. (1985). Locke and the innatists. *History of Philosophy Quarterly*, *2*(4), 411–420.