BRAIN-MACHINE INTERFACE COUPLED COGNITIVE SENSORY FUSION WITH A KOHONEN AND RESERVOIR COMPUTING SCHEME

BY

ONYEAMA OSUAGWU

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor Stephen Levinson, Chair
Associate Professor Lynford L. Goddard, Chair
Professor P. Scott Carney
Professor Alfred Hübler
Assistant Professor Maxim Raginsky

ABSTRACT

Artificial Intelligence (AI) has been a source of great intrigue and has spawned many questions regarding the human condition and the core of what it means to be a sentient entity. The field has bifurcated into so-called “weak” and “strong” artificial intelligence. In weak artificial intelligence reside the forms of automation and data mining that we interact with on a daily basis. Strong artificial intelligence can be best defined as a “synthetic” being with cognitive abilities and the capacity for presence of mind that we would normally associate with humankind. We feel that this distinction is misguided. First, we begin with the statement that intelligence lies on a spectrum, even in artificial systems. The fact that our systems currently can be considered weak artificial intelligence does not preclude our ability to develop an understanding that can lead us to more complex behavior. In this research, we utilized neural feedback via electroencephalogram (EEG) data to develop an emotional landscape for linguistic interaction via the android's sensory fields which we consider to be part and parcel of embodied cognition. We have also given the iCub child android the instinct to babble the words it has learned. This is a skill that we leveraged for low-level linguistic acquisition in the latter part of this research, the slightly stronger artificial intelligence goal. This research is motivated by two main questions regarding intelligence: Is intelligence an emergent phenomenon? And, if so, can multi-modal sensory information and a term coined called “co-intelligence” which is a shared sensory experience via coupling EEG input, assist in the development of representations in the mind that we colloquially refer to as language? Given that it is not reasonable to program all of the activities needed to foster intelligence in artificial systems, our hope is that these types of forays will set the stage for further development of stronger artificial intelligence constructs. We have incorporated self-organizing processes - i.e. Kohonen maps, hidden Markov models for the speech, language development and emotional information via neural data - to help lay the substrate for emergence. Next, homage is given to the central and unique role played in intellectual study by language. We have also developed rudimentary associative memory for the iCub that is derived from the aforementioned sensory input that was collected. We formalized this process only as needed, but that is based on the

assumption that mind, brain and language can be represented using the mathematics and logic of the day without contradiction. We have some reservations regarding this statement, but unfortunately a proof is a task beyond the scope of this Ph.D. Finally, this data from the coupling of the EEG and the other sensory modes of embodied cognition is used to interact with a reservoir computing recurrent neural network in an attempt to produce simple language interaction, e.g. babbling, from the child android.

DEDICATION

To my parents, Chief Dandison Osuagwu-Akpunku and Princess Lolo Joyce Osuagwu-Akpunku, for their never ending love and support. You have always been my role models, I am forever grateful for your sacrifice in coming to a new and foreign land with the hope of making a better future for your children. Know that this was not in vain. May you both rest in eternal peace. To all those who came before me, to closed doors and hostile spaces: I thank you from the bottom of my heart for your sacrifice and I promise not to waste it. For those who have supported this eccentric character in search of higher learning and a purpose in life, thank you. Until we meet again...

Amaja -ịbụ maara ihe, ka anyị wee na-aga kpakpando na karịrị - Dare enim se esse sapientes sic itur ad astra praeter

ACKNOWLEDGMENTS

I would first like to thank my advisers, Professor Stephen Levinson and Professor Lynford Goddard. I have been truly blessed to be allowed to be a part of your groups. I couldn't have asked for better, more appropriate advisement if I had designed this Ph.D. program myself. Getting the best of both worlds in academia due to the different experiences you both bring to the discussion has broadened my philosophy in ways this simple acknowledgment can only begin to capture my gratitude to you both.

Professor Goddard, you have truly shown me the virtue of patience and forgiveness in research. I had many ideas when I first joined your lab and I wanted to work on all of them. Although this stunted my productivity, I later realized this was the way to discovering who I was as a scientist, mathematician and an engineer. None of this would have been remotely possible without your continued and intentional support. For this I am forever grateful.

Professor Levinson, you allowed me to join your lab in my fourth year which is why I refer to my time in your lab as my first and only post-doc. I was able to move in a totally different research direction, as a result. Your sage advice about life, research and career have served me well during many talks and presentations. You also gave voice to an alternative path for the Ph.D. in industry which is rarely spoken of with such personal knowledge in the world of academia. I was able to learn and truly appreciate the many, many components that influence this phenomenon we would like to call intelligence. I just hope that I can make a contribution of merit as a result of such stellar tutelage.

To my committee members, Professor Carney, Professor Hübler and Professor Raginisky, I have long been an admirer of all of your individual research efforts, thank you for accepting the call and honoring me by joining me on this journey. Thank you for the advice during our many chance encounters on campus and meetings to discuss an idea in greater detail.

I would like to thank the members of the Language Acquisition and Robotics Research Group. It can be said without any hesitation that we have a rather interesting and eclectic collection of researchers. The debates and discussions that I was present for and the arguments that sometimes ensued were well worth the information gleaned. Researchers should always be ready to stand alone in their convictions in an idea. Thank you for letting me know that I can do exactly that when called upon. In particular, I would like to thank Dr. Alexander Duda, Dr. Lydia Majure, Jacob Bryan and Luke Wendt for very interesting conversations about math, intelligence and the like. My two former undergrads Junia Findlay and Sameer Bojja are due for praise, Junia in particular for a top-notch job on a very difficult problem with very little supervision. Working with you guys has been great.

I would like to thank all of the members of the Photonics Systems Laboratory for enduring years of presentations on my attempts at neuromorphic optical computing that although possibly interesting was not of direct relation to the work that many members where engaged in. It has been a real thrill these many years.

Thank you to Professor Philip Phillips who I first met at a physics conference many years ago and was the person who helped to recruit me to attend this fine institution and called me to let me know I was accepted in the Physics Ph.D. program. Thank you to Professor Pat Ransom for being the first person to recognize that I was more interested in electrical and computer engineering than I let on. As a result, I switched departments and this has been the singular best decision I have made in my academic career. No slight towards the physics department which is great. I would also like to thank Kathy Reeves of the SURGE fellows office for her support. Thank you to Professor Kay Kirkpatrick for many spirited discussions on mathematics, the world and critical presentation skills for success. Your algorithm for job interviews to offers ratio was spot on! Thank you to Professor Kimani Toussaint, whose sage advice helped tremendously when exploring the job market and opportunities in general. Thanks to S. Lance Cooper for great advice and recommendations during my time in the Physics Ph.D. program at Illinois.

Special thanks to Toni Pitts and Rebecca Wiltfong, you guys have been one of the most stable points of friendship and support during my time here at Illinois. This experience truly wouldn't have been the same without you. The ECE office was always a great support to me during my stay as well. Jen, Laurie, Terry and the list goes on! Thank you for being there. To all of my friends in the MNTL Business office, thank you so very much for treating me like one of the family and providing so much unexpected support! Kelly, Tasha and Nandini you guys rock! (This goes for the new hires and the temps that left as well!)

My graduate student friends, where do I begin? We have commiserated so many times over the years, but now it is time for the next chapter in our lives. Yemaya Bordain, first female African-American ECE Ph.D. and my friend. Now that it is over, I can already say that it was worth it. Ritoban Basu-Thakur, my brother from another mother as one would say. Thank you for your support and friendship from the Physics Ph.D. to the end. Alex Weiss, I think we have tried everything on the menu at Lai Lai Wok, it must definitely be

time to go! Drs. Jeremy Tan, Georgios Tsekenis, Sharnali Islam and Maksim Sipos, thank you all for your friendship and support. I would like to thank Cliston Cole for his support, it is your turn to defend now! I'll be waiting for you at the job. I want thank all of my friends in the organization I co-founded, Graduate Engineers Diversifying Illinois (GEDI), which wouldn't have happened if not for the vision of Dr. Phoebe Lenear, the Mother of GEDI and Professor Andrew Alleyne. Shaneen, Lonna, Yemi and Desiree, I am sure you will take it to new heights!

Finally but definitely not least, I would like to thank my family, both in blood and extended, for all of your support and encouragement throughout this process. You guys believed in me even when, sometimes, I didn't believe in myself. Thank you to my dearest Deanna, Ibe, Chima, Tia, Paulie, Alisa and Mr. and Mrs. Martin for all of your encouragement. In particular, my big sister Ure, thank you for the hours of advice and counseling on the phone when I hit some really rough patches in life while I was out here in the corn fields.

For all those who know me, we are well aware that I know far too many people to list all of their names in this document. Let it suffice to say that I am grateful for the way that each and everyone of you have enhanced my time here at the University of Illinois at Urbana-Champaign.

Lastly, I would like to thank God. All things are possible through you. I thank you everyday for allowing me the many opportunities to share these blessings and help to make this world a better place.

CONTENTS

LIST OF ABBREVIATIONS

ASIC	Application Specific Integrated Circuit
AI	Artificial Intelligence
ANN	Artificial Neural Network
BMU	Best Matching Unit
LAR	Language Acquisition and Robotics Group
LSN	Liquid State Machine
ESN	Echo State Network
EEG	Electroencephalogram
FPGA	Field Programmable Gate Array
FAP	Fixed Action Pattern
f-MRI	Functional Magnetic Resonance Imaging
HMM	Hidden Markov Model
NN	Neural Network
SOM	Self-Organizing Map
SiN	Silicon Neuron
SNN	Spiking Neural Network
RNN	Recurrent Neural Network
RC	Reservoir Computing

Chapter 1

INTRODUCTION

Without language, thought is a vague, uncharted nebula. – Ferdinand de Saussure

Language is the light of the mind. – John Stuart Mill

We have taken for granted one of the most important components of intelligence our species has developed, language. We purport to study intelligence in artificial systems and during such journeys we have looked into many versions of systems that display intelligence in an attempt to find and replicate that spark of life in research. It remains an elusive goal, but these chapters discuss our latest attempts toward a smart system in the form of an android named Bert.

In our original approach, we leaned heavily on an instinctual framework because in many ways it can be considered a catalyst for many more complex intelligent behaviors. Without instincts and drives, it is doubtful that many of the expressions of intelligence we observe would ever have arisen from our evolutionary predecessors. This tact, unfortunately, led us away from a more fundamental assumption that has driven this work, "functional equivalence". As a result of this refocusing, this dissertation has landed squarely on the language components of the research. Language provides the scaffolding for the deluge of information and patterns that appear around us. It also provides us with the tools to categorize and recognize our experiences and those of others providing us with the facility to manipulate and evolve these symbols into thoughts that we can express on a grander scale. We also advocate for embodiment in this research. The majority of the concepts we express and experience in language are through our physical forms. This undoubtedly colors our collective understanding and helps to shape the language we use and develop.

In concrete terms, this dissertation is about very simple natural language acquisition in particular, more specifically, attempting to empower our android, Bert, to have a basic ability to attempt to learn words and "contextualize" simple components of language through visual, auditory and emotional senses. We have incorporated brain-machine interfaces via electroencephalogram and multiple sensory input as embodied cognition concepts in this paradigm because they provide closure for the feedback loop and memory components of this work.

This research is not a Messianic quest to deliver true artificial intelligence, cognition or the like. These types of quests tend to be dangerous as they usually require a dogmatic devotion to a technique or philosophy. Admittedly, it is possible that we may fall short of some of these goals because of the sheer number of coordinated efforts that must be brought to bear just to even attempt this work within a team, much more so in a single Ph.D. dissertation performed by a single graduate student. At this stage in the study, we would like to keep the options open for the big problems of true "strong" artificial intelligence.

In the pursuit of understanding intelligence, and hopefully cognition and consciousness, mathematics is the language of choice to represent the world in physics and engineering. This can be both a benefit and a burden. We have reaped untold rewards from routinely adhering to the tenets of this language, and as a result, we are inclined to believe that as it is constructed its powers are without limit. The first person, to our knowledge, to question this successfully was Kurt Gödel in "On Formally Undecidable Propositions of Principia Mathematica and Related Systems" [1]. It is believed that a pure definition of human intelligence in the more rigorous traditions of mathematics does not exist in the literature, although there have been some attempts to define this for machines [2, 3, 4, 5]. We will provide several attempts at a definition from several groups and then offer what are considered to be the most salient features and a construction of intelligence that will be the mainstay of our theoretical and experimental approach.

From "Mainstream Science on Intelligence" [6], an editorial statement by fifty-two researchers:

> *A very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surrounding "catching on," "making sense" of things, or "figuring out" what to do.*

"Intelligence: Knowns and Unknowns" [7], a report published by the Board of Scientific Affairs of the American Psychological Association states the following:

> *Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to*

> *engage in various forms of reasoning, to overcome obstacles by taking thought. Although these individual differences can be substantial, they are never entirely consistent: a given person's intellectual performance will vary on different occasions, in different domains, as judged by different criteria. Concepts of "intelligence" are attempts to clarify and organize this complex set of phenomena. Although considerable clarity has been achieved in some areas, no such conceptualization has yet answered all the important questions, and none commands universal assent. Indeed, when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen, somewhat different, definitions.*

We take issue with the bifurcation within the field of artificial intelligence. This can be demonstrated by what we would like to call the "path problem". Given a path (synonymous with a problem or information in this definition) to follow, an increasing number of our supposedly "intelligent" systems can follow them with little to no supervision. Now this would seem to be exactly the kind of behavior we are looking for in an "intelligent" agent, except for a major component of biological "real" intelligence that is obscured by this slight of hand, another real intelligence had to determine and present the path that this agent is following. Even in the cases in which an agent can identify a path on its own, this is still driven and limited by its programming. A truly intelligent agent would have the ability to identify, define and solve a problem or learn a task independent of another intelligence, i.e. humanity. A simple example is a calculator. We can use a calculator to obtain the answer to many arithmetic problems, but calculators currently do not solve problems on their own or try to understand the process and result of the calculation. Spoken more directly, automation is not cognition! This is the difference between how we do things and why we do things.

Realistically this "problem" we just described is called weak artificial intelligence (AI) also known as weak AI in general and the "Chinese Room" in particular and it might not be such a problem at all. All of the tests we use for intelligence shows that it develops on a continuum and not in discrete states. As a result, we think that it is naive to assume that a large body of AI research appears to have assumed the latter. The few cases we could reliably call weak AI are experiencing a boom like never before. This has marched hand in hand with our ability to collect data from a myriad of locations, giving us the term big data. As computational resources continue to drop in price and increase in power and availability, we have been able to tackle many problems that benefit from this amalgamation of analytical techniques.

This dissertation will venture into the philosophical and computational aspects of a simple intelligent activity, instinct, with the hope of developing a framework that will lead to conceptualization and memory formation. This will play in to simple language acquisition and semantics. It is thought that if intelligence is to be studied in the way that humans generally appreciate it, semantics must play a significant role. This understanding may also help illuminate why our attempts at artificial intelligence have borne little fruit. We begin with a discussion of the brain and mind in the biological sense and hopefully explain to the reader why we should not be beholden to this description.

1.1 The Brain, Mind and Language

The mind has been a great source of intellectual fodder for most of human existence. These discussions have ventured into realms as varied as the supernatural to religion and science. In the book *Mathematical Models for Speech Technology,* Levinson [8] provides an excellent table referring to the implementation of a particular era that was used to describe the mind. Table 1.1 appears in [8].

Table 1.1: Mental Representations Throughout History

Pre-Industrial Period	Industrial Period	Information Age	Mathematical Abstraction
Rudder	Governor Thermostat	Feedback Amplifier	Control Theory
Hydraulic Systems	Telegraph	Internet	Communication Information Theory
Wax Tablets	Photographic Plates	AUDREY	Pattern Classification Theory
Clocks	Analytic Engine	ENIAC	Theory of Computation

We currently do not have a well-defined mathematical conceptualization of the mind despite our many in-depth images and diagrams of the brains of many species. Admittedly, this might be because of our fixation on the organ, but it is the most logical starting place for this kind of work. This is also the heart of Section 1.1.1.

1.1.1 The Neuroscientific Approach

The field of neuroscience deals with the structure and function of the nervous system and brain. As a result the approach to intelligence and memory extends from biological origins.

In this work, we are more concerned with cognitive neuroscience which investigates the biological substrates involved in cognition, with a specific focus on the neural substrates of mental processes [9]. Ultimately, the goal of the neuroscientist is to understand the underlying biological systems that allow the brain to function and drive all of its functions including but not isolated to intelligence.

1.1.2 Artificial Intelligence Approach

Artificial intelligence (AI) is intelligence exhibited by machines and/or software. Major AI researchers and textbooks define AI as “the study and design of intelligent agents” [5**?**], where an intelligent agent refers to a system that can perceive its environment and takes actions to achieve a goal. The term originated from John McCarthy in 1955 who defines it as "the science and engineering of making intelligent machines" in an online discussion on the topic. This is where my research is rooted and this is where most of the divergence from the neuroscientific methods develop, in particular with the notion of “functional equivalence”.

1.2 Language

Language is center stage in this work. It is the most important tool that intelligent systems use. It can be argued that in looking at language we are focusing on the shadow of intelligence and not intelligence itself. There is certainly truth to this statement, but in order to quantify intellectual capability in the android used in this dissertation, It was necessary to choose a modalilty, and language is by far the most meaningful to human interactions. When we discuss language we really are concerned with words or symbols, the methods of combining them as a systematic means of communicating ideas or feelings.

1.3 Emotions

Emotions also play a role in the learning of languages [10, 11, 12]. This aspect of the human condition is often left out of studies of intelligence due to difficulty in the quantification of states. In order to capture a more accurate picture of a word or a string of words, it would be hard to construct simple mental states in silicon without access to this information. In this research, we use a brain-machine interface to provide access to emotional data for exactly that purpose.

1.4 Cybernetics

The term cybernetics was originally coined by Norbert Wiener in his original book *Cybernetics, or Control and Communication in the Animal and the Machine* [13]. Cybernetics is a cross-disciplinary approach for exploring regulatory systems, also known as homeostasis by Wiener, their structures, constraints, and possibilities. Cybernetics was developed to help elucidate the study of systems, mechanical, physical, biological, cognitive, and social systems, for example. Cybernetics is applicable when a system being analyzed generally possesses some type of feedback. By this we mean that an action by the system generates some change in its environment and that change is reflected in that system in some way that causes the system (or another coupled system) to change. As this work utilizes synthetic instincts, many of the principles from cybernetics will be needed.

1.4.1 Functional Equivalence

Alan Turing made an argument that in order to study the mind one does not necessarily need to have a full understanding of its inner workings. All that would be required is to faithfully reproduce its functionality [14, 15]. This belief is central to most artificial intelligence research and provides us with a freedom to develop systems unconstrained by our lack of understanding of the biological brain and the mind that emerges from it. To this end, the dissertation focuses on developing and researching technology that can reproduce some of the basic properties of systems that we can regard as residing on a scale of intelligence. We intend to leverage this in order to classify and/or identify possible intellectual states that truly do not require a faithful representation of a biological brain in any way. We take this concept to its logical end by developing actions without direct "neuronal" counterparts that can drive similar actions. The best analogy for this way of approaching the problem is reflected in the time line of physics research. Newtonian mechanics and electricity and magnetism developed well before we had any true measurement or understanding of the atom and electron. This did not preclude scientists from conducting experiments on the properties that emerged from these particles and their interactions. When dealing with emergent properties it is always useful to consider the scale of detail and description required for the individual units that we are coaxing to reveal their intrinsic properties.

1.5 The iCub Android Platform

In the laboratory we have an iCub humanoid robot which we affectionately call Bert in Figure 1.1. Bert is an advanced replica of a four to five year old child and a great tool for the study of embodied cognition. He is 3.5 feet tall and weighs 48 pounds. Bert also has 53 degrees of freedom in joints, two eye-cameras (30 Hz), ear-microphones with external pinnae for localization, proprioception (joint position, speed, and torque), and a head-gyroscope for orientation.

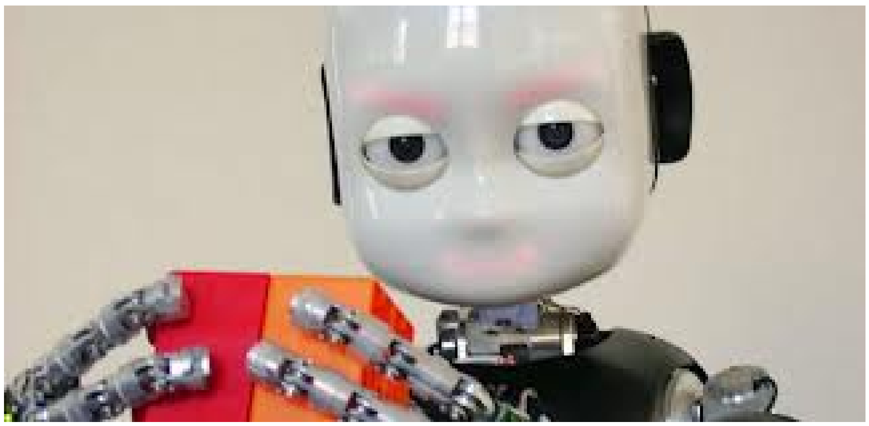


Figure 1.1: The iCub android.

1.6 "Weak" AI, "Strong" AI and the Intelligence Continuum

After studying this problem, we have come to the realization that this segmentation between weak and strong AI is a bit loaded. Natural intelligence occurs on a continuum and, as a result, discounting the levels of intellectual performance can actually hinder our understanding of this phenomenon. Intelligence, in the creatures that we agree have this capability, usually starts in a state that could be considered weak and becomes strong from interaction and experience.

1.6.1 Machine Learning

Machine learning is by no means the only version of artificially intelligent technologies utilized by the research public, but it has gained quite a lot of attention in the public eye due to everyday interactions with computer agents such as Apple's Siri, Google's Google Now service and the soon to be released Microsoft product, Cortana. Machine learning is essentially a collection of techniques for solving a small domain of problems with limited to no supervision. It is essential to work in these domains at this stage, but we have to remain ever mindful of the larger goal of artificial intelligence.

1.6.2 To Serve Humankind or to Understand Humankind

When dealing with the development of artificial intelligence, we are forced to ask the question "Whom will it serve?" We feel that this question is misguided by our Western paranoia regarding invasion and conquest. But unfortunately, this question has shaped the type of research that is conducted in this area. If we wish to develop automatons that will follow our instruction with little guidance we are obviously already there. Our cell phones can already do quite a lot of these things without any of us having to lift a finger. The deeper questions that we should be asking do not lie in the serving of humankind, but in the understanding of humankind, in particular, that is, what does it mean to be sentient? This is a question for the ages, of course, and we make no allusions to providing an answer in this dissertation, but we will say that this question has played a profound role in our decision to pursue this research.

Chapter 2

A REVIEW OF RELEVANT LITERATURE

If you wish to learn to dance, you should watch the feet of others. –African Proverb

2.1 Instinctual Behavior

Instinct or innate behavior is the inherent inclination of an organism to demonstrate or participate in a particular complex behavior. A simple instinctive behavior is a Fixed Action Pattern (FAP). FAPs are composed of very short to medium length sequences of actions, without variation, that are carried out in response to very specifically defined stimulus.

If a behavior is performed without being based upon prior experience, in the absence of learning, that behavior is instinctive and is therefore an expression of innate biological factors. Honeybees communicate by dancing in the direction of a food source without formal instruction. Sea turtles, newly hatched on a beach, will automatically move toward the ocean. Other examples include animal fighting, animal courtship behavior, internal escape functions, and the building of nests. All of these are examples of complex behaviors that are not learned or trained prior to the necessity to use that action.

An instinct is not to be confused with the reflexes, a simple response of an organism to a specific stimulus, such as the contraction of the pupil in response to bright light or muscle spasming when the knee is tapped. Instincts are inborn complex patterns of behavior that exist across a majority of a species and in several cases even across species. The absence of volitional capacity does not mean that a species has an inability to modify fixed action patterns. For example, people may be able to modify a stimulated fixed action pattern by consciously recognizing the point of its activation and simply stop doing it, whereas animals without a sufficiently strong volitional capacity (also known colloquially as self-control and awareness) may not be able to stop their fixed action patterns, once activated [16].

In the quest to observe intelligent behavior in inanimate systems we should be mindful of the old saying within the acting profession: “What’s my motivation?” In the case of this research, the motivation is the first step in attentive behavior that will help us to construct very simple concepts from verbal babbling during language acquisition.

2.2 Brain-Machine Interfaces

This technology allows data collection from a brain, human or otherwise, and processed with a computer [17]. It is a tool of importance in neuroscientific endeavors and has several incarnations that have been used to control robotic limbs and generate speech via focusing on letter clusters on a monitor for people with physical limitations.

2.2.1 Electroencephalogram (EEG) and Language

Electroencephalogram (EEG) technology had its start in 1875, when a physician named Richard Caton began to investigate the possibility that electric signals might be detected from exposed cerebral hemispheres of rabbits and monkeys [18]. Its use in language occurred much later, around 1975 [19, 20], in the form of Event-Related Brain Potential (ERP) measurements.

2.3 Neural Networks

Neural networks are an intriguing area of study and variations on these designs have played a role in this research. The literature generally falls into two categories: biological neural networks or artificial neural networks. We begin with a discussion about biological neural networks.

2.3.1 Biological Neural Networks

Biological neural networks, e.g. biological neurons, synapses, etc. are under intense and active investigation. A large portion of this work specifically focused on identifying the regions of the brain, the human brain in particular, and their functions. Techniques such as Electroencephalogram (EEG) and Functional Magnetic Resonance Imaging (f-MRI) are valuable tools for the study of in-vivo brain activity, but unfortunately we have been hindered by our current resolution issues regarding in-vivo neuronal activity. f-MRI is a powerful tool [21, 22], but uses the correlation between blood deoxygenation to relate information and activity within the brain to the observer. This also has the very clear side effect of a significant time delay between neuronal activity and readout from the f-MRI. f-MRIs do however provide us with the deepest in-vivo brain information to date. The EEG reports actual electrical activity but it is limited in that this information is gathered as an average over regions of the brain and is generally only reliably from the cortical neurons. This coupled with the inverse problem which stops us from directly identifying the neurons associated with the signals we are recording can be problematic to say the least. EEGs main contribution to the advancement of the study of the brain is its real-time information which is directly correlated to neuronal activity, not blood deoxygenation. There still remain issues regarding the interpretation of the data collected by this technique and its predictive value [23]. The main goal of all the biological neural network studies is to, in effect, reverse engineer the human brain. Recent inroads toward this goal include the human connectome project with its goal of generating a complete map of the human brain via MRI technology and mapping the conductive flow of blood through this organ [24, 25, 26]. This goal has remained elusive for quite some time and will most likely remain this way well into the future. Admittedly having a full picture of a human brain or several brains with the fiber bundles in tact, does not promise full understanding of the system in its own right, but it is considered one of many uses of new technology to gather more information on the topic. Surely we would be remiss if we did not mention the pioneering work of Hodgkin and Huxley [27] on the squid neuron which inspired quite a bit of modern research by providing a mathematical relationship for the voltage changes within a neuron or "spiking" as it is more commonly called.

2.3.2 Artificial Neural Networks

Artificial neural networks, e.g. Hopfield networks, feed-forward networks, etc. began with the work of McCulloch and Pitts in their seminal work, [28]. In this paper, the concept of "threshold logic" was introduced where stimulus would be allowed to accumulate and would only trigger a response from the system once a specified threshold was achieved. This was, of course, what was seen in several biological instances of neural activity and as a result considered a reasonable first-order model. We know that the specifics of neural activity are more complicated than initially laid out in this paper, but the idea has had a resurgence in recent years with the development of faster computing technology capable of supporting the intensive resources necessary for large-scale simulation and the solution of the exclusive-or circuit by Werbos in 1974 [29]. These systems needed to be tested in a practical manner in order to verify that there was anything of merit to this framework. To this end, many of these systems have proven themselves very adept at discovering and/or recognizing certain patterns in data. Supervised neural networks can display this kind of pattern recognition with a specified amount of pre-training. Unsupervised neural network models had their start with the work of Hebb and his hypothesis for neural plasticity generally referred to as "Hebbian learning" [?]. Two paradigms in neural networks research utilized during this study are discussed in section 2.3.2.1.

2.3.2.1 Self-Organizing Maps

Self-Organizing Maps (SOM) also known as Kohonen maps were first developed by Teuvo Kohonen [30]. They are a class of artificial neural networks that are particularly useful for finding patterns in data that are gathered without any pre-classification. The training is unsupervised and with the right set of parameters, the network generates results that are useful for finding associations or groupings in data. This method also has the benefit of being an unsupervised technique, which dovetails with this research by allowing passive clustering of sensory inputs. Analyzing large datasets with multiple variables in each sample is generally considered a complicated task since it is difficult to visualize the similarities between samples in greater than two or three dimensions. A SOM serves to reduce the dimensionality of a problem by mapping these multiple factors onto a two-dimensional space or grid. The distance between nodes on a grid not only demonstrates which samples are similar but also gives us a sense of the topology of the data in the space we are investigating. A diagram for the self-organizing map can be seen in Figure 2.1.

The basic Kohonen map algorithm goes as follows:

1. Initialize the weights to small random values.
2. Select an input vector from the data: x_i.
3. Find the Best Matching Unit (BMU), i.e. the node that has weight values closest to the input vector. Typically, this is done by calculating the Euclidean distance for each node and finding the minimum node.
4. Adjust nodes in the neighborhood of the BMU so that their weights are closer to the input.
5. Adjust weights according to the following equation:

$$W(s+1) = W(s) + \Theta(i,s)\alpha(s)(X(s) - W(s)) \tag{2.1}$$

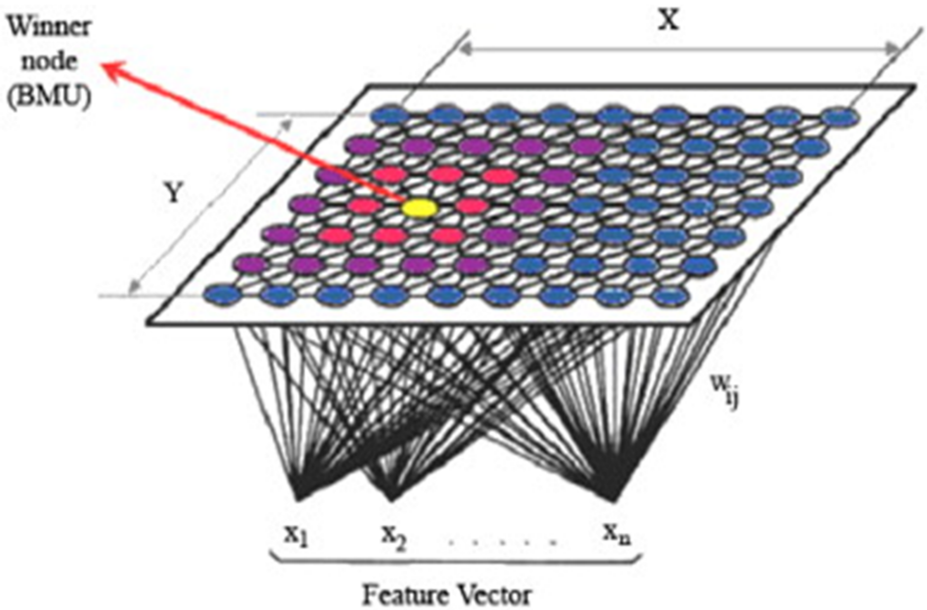


Figure 2.1: 2D self-organizing map. Nodes with the greatest similarity to the winning node are updated to with similar weights.

Θ is the neighborhood function that decreases in radius over time,α is the learning rate that also decays over each epoch, $X(s)$ is the input vector, and s is the epoch count for the algorithm.

6. Present the samples of each epoch (cycle) in a different order until the total error decreases below a specified threshold.

Self-organizing maps have been suggested as a possible description of biological neural activity with spike timing dependent plasticity [31]. There has also been work in this domain directed toward applications in speech, i.e. Spatio-Temporal Organization Maps (STOM) and the like [32, 33, 34], and have also been explored with regard to memory formation [35] which are both of particular interest to this work.

2.3.2.2 Reservoir Computing

Reservoir Computing (RC) is the collective name given to techniques and formulations of Recurrent artificial Neural Networks (RNN) that stem from temporal recurrent neural networks [36], liquid state machines [37], echo state networks [38], and decorrelation-backpropagation learning [39]. This model has been suggested as an elementary framework for how learning and general intelligence occurs in the human brain [40].

2.3.2.3 Temporal Recurrent Networks

What we call reservoir computing owes a significant part of its history to the work of Dominey. In particular, his research on cortico-striatal circuits in the human brain (e.g., [41, 42], and beyond). His work in cognitive neuroscience and functional neuroanatomy with the expressed goal of modeling and studying complex neural structures over theoretical computational work guided his focus on these particular models of learning and computation. Dominey also discusses a major component of the reservoir computing model as follows: "... there is no learning in the recurrent connections, only between the State units and the Output units. Second, adaptation is based on a simple associative learning mechanism ..." [43]. Dominey also goes on to talk about the neural reservoir module as a temporal recurrent network which requires randomization in the reservoir connections.

2.3.2.4 Echo State Networks

Echo State Networks (ESNs), see Figure 2.2, are an architecture and a set of supervised learning principles for recurrent neural networks. The main idea is (i) to drive a random, large, fixed recurrent neural network with the input signal, thereby inducing in each neuron within this "reservoir" network a nonlinear response signal, and (ii) combine a desired output signal by a trainable linear combination of all of these response signals. ESNs represent one of the two pioneering RC methods. The approach requires that a random recurrent neural network possess certain algebraic properties, training only a linear readout from it is often sufficient to achieve excellent performance in practical applications. The untrained RNN part of an ESN is called a dynamical reservoir, and its states are termed echoes of its input history. ESNs, as a standard, use "weighted sum and nonlinearity" type of simulated analog-valued neurons, most often with a tanh nonlinearity. Leaky integration of the neurons' state recently became a standard practice in ESNs [44]. Classical recipes and conditions of producing the ESN reservoir were outlined in the original introduction of ESNs. The readout from the reservoir is usually linear. The original and most popular batch training method to compute the output weights is linear regression. For online training settings the computationally cheap least mean squares algorithm is recommended [45]. Some of the first ESN publications were framed in settings of machine learning and nonlinear signal processing applications. The original theoretical contributions of early ESN research concerned algebraic properties of the reservoir that make this approach work in the first place and analytical results characterizing the dynamical short-term memory capacity [46] of reservoirs.

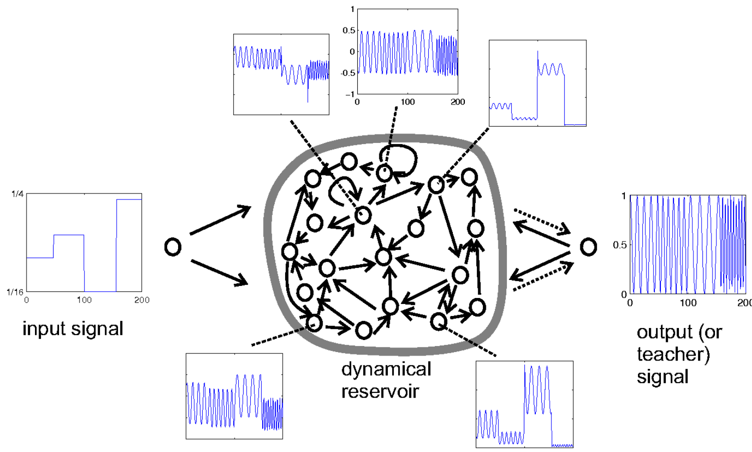


Figure 2.2: Echo State Network. Adapted from a presentation by Herbert Jaeger [46].

2.3.2.5 Liquid State Machines

Liquid State Machines (LSMs) represent the other pioneering reservoir method, developed independently from and simultaneously with ESNs. LSMs were developed from a computational neuroscience background, aiming at elucidating the principal computational properties of neural microcircuits [47, 48, 49]. Thus LSMs use more sophisticated and biologically realistic models of spiking integrate-and-fire neurons and dynamic synaptic connection models in the reservoir. The connectivity among the neurons often follows topological and metric constraints that are biologically motivated. In the LSM literature, the reservoir is often referred to as the liquid, following an intuitive metaphor of the excited states as ripples on the surface of a pool of water. Inputs to LSMs also usually consist of spike trains. In their readouts LSMs originally used multilayer feed-forward neural networks (of either spiking or sigmoid neurons), or linear readouts similar to ESNs. Additional mechanisms for averaging spike trains to get real-valued outputs are often employed. RNNs of the LSM-type with spiking neurons and more sophisticated synaptic models are usually more difficult to implement, to correctly set up and tune, and typically more expensive to emulate on digital computers (with a possible exception of event-driven spiking Neural Network (NN) simulations, where the computational load varies depending on the amount of activity in the NN) than simple "weighted sum and nonlinearity" RNNs. Thus they are less widespread for engineering applications of RNNs than ESNs. However, while the ESN-type neurons only emulate mean firing rates of biological neurons, spiking neurons are able to perform more complicated information processing, due to the time coding of the information in their signals (i.e., the exact timing of each firing also matters). Also findings on various mechanisms in natural neural circuits are more easily transferable to these more biologically realistic models. The LSM version of RC consists of analytical characterizations of the computational power of such systems [50].

2.3.2.6 Backpropagation-Decorrelation

The idea of separation between a reservoir and a readout function has also been arrived at from the point of view of optimizing the performance of the classical RNN training algorithms that use error backpropagation. In an analysis of the weight dynamics of an RNN trained using the Atya-Parlos recurrent learning (APRL) algorithm [51], it was revealed that the output weights of the network being trained change quickly, while the hidden weights change slowly and in the case of a single output the changes are column-wise coupled. APRL effectively decouples the RNN into a quickly adapting output and a slowly adapting reservoir. These results lead to a new iterative/online RNN training method, called BackPropagation-DeCorrelation (BPDC). It approximates and significantly simplifies the APRL method, and applies it only to the output weights, turning it into an online RC method. BPDC uses the same type of neurons as ESNs. BPDC learning is claimed to be insensitive to the parameters of the fixed reservoir. BPDC boasts fast convergence times and thus is capable of tracking quickly changing signals.

2.4 Swarm Intelligence

Swarm intelligence deals with natural and artificial systems composed of multiple agents that coordinate using self-organization and decentralized control. The discipline has focused on the collective behaviors that emerge from local interactions of these agents within group and externally with their environment. Ant colonies and termites, schools of fish, flocks of birds, herds of land animals are all good examples of the types of systems studied by swarm intelligence [52]. Human behavior can also be studies in a swarm intelligence fashion, mob behavior and multi-robot systems. Swarm intelligence requires an interdisciplinary approach to study systems in such a wide variety of domains.

It is customary to divide swarm intelligence research into two areas according to the nature of the systems under analysis. We speak therefore of natural swarm intelligence research, where biological systems are studied; and of artificial swarm intelligence, where human-like behaviors are studied.

A swarm intelligence system has the following properties:

- It is composed of a large number of individuals or agents.

- The agents are relatively homogeneous (i.e., they are either all identical or have very little variation).
- The interactions among the agents are based on relatively simple behavioral rules that exploit only local information that the agents exchange directly or via the environment.
- The overall behavior of the system results from the interactions of agents with each other and with their environment, that is, the group behavior self-organizes.

Swarm intelligence systems are most known for their ability to act in a coordinated way without the presence of a coordinator or of an external controller. In nature we observe swarms that perform some collective behavior without any agent controlling the group, or with no mechanism of global awareness of the overall group behavior. This still does not preclude the swarm as a whole from showing intelligent behavior. The interaction of a spatially neighboring agents that follow simple rules has been the main explanation for such interesting and complex behavioral patterns.

It helps to describe this behavior of each agent in the swarm in probabilistic terms: Each agent displays stochastic behavior that depends on its local access to data within the neighborhood. This allows us to design swarm intelligence systems that are scalable, parallel, and fault tolerant.

Scalability means that a system can maintain its function while increasing its size without the need to redefine the way its agents interact. A swarm intelligence system interactions involve only neighboring agents, the number of interactions usually do not grow with the overall number of agents in the swarm: each agent's behavior is only loosely influenced by the swarm dimension. In the artificial system studied for this work, scalability is useful because a scalable system can increase its performance by simply increasing its size, without the need for additional coding.

Parallel action is possible in swarm intelligence systems because the agents composing the swarm can perform different actions in different places at the same time. Again in this study, parallel action can be helpful because it can make the system it helps to foster the ability to self-organize in teams that take care of different tasks.

Fault tolerance is an inherent property of swarm intelligence systems since it is decentralized, self-organized in the nature of their structures. Because the system is composed of many interchangeable agents and none of them is in charge of controlling the overall system behavior, a failing agent can be easily removed and replaced by another.

2.5 Statistical Mechanics, Emergent Phenomena and Phase Transitions

In nature, there are many examples of systems that exhibit properties en masse that cannot be discerned from the individual parts. One of the best examples of this being the wetness of water. Individual molecules do not impart the sensation of wetting, but after a certain number of them are placed together at the right temperature we can sense this property. There have been similar arguments made regarding the brain and neurons as a dynamical system. This property known as emergence has Hopfield-related artificial neural activity also seen in spin glasses [53] a disordered magnet, where the magnetic spins of the component atoms (the orientation of the north and south magnetic poles in three-dimensional space) are not aligned in a regular pattern. These types of systems have basic similarity to the brain and could yield some mechanisms for the development of neural states. In connection to this body of work, this contributes some mechanisms for observing and classifying emergence during the experiments.

In the case of cellular automata or simple agents [54, 55, 56, 57, 58, 59, 60] several papers have explored the properties of these systems through the lens of statistical mechanics. The swarm in this case plays the role of adaptation for the concepts in this model.

2.6 Natural Language Processing

The ability to communicate, transcribe, represent ideas, situations, objects and the relationships between them both real and imagined can quite literally be seen as one of the greatest hallmarks of human intelligence, without equal. So quite naturally in the world of artificial intelligence we are keenly interested in getting our systems to understand and produce language in a natural way, indistinguishable from humans. There are many facets to language as studied by linguists, anthropologists, sociologists, scientists and engineers. Some of these components are listed below which are clearly relevant to the Lengua Franca of the United States, English (pun intended).

2.6.1 Syntax and Production Rules

Syntax provides the rules for order, principles and process of writing and speaking grammatically correct sentences. This allows us to have some level of uniformity when communicating ideas which is a large part of language's purpose. In the early stages of language acquisition, we struggle with this component since we can easily name objects and actions without following these prescribed rules. This did not play any role in the research in this dissertation. Production rules are the simple rules for the generation of a syntactically correct piece of grammar. They come in several varieties that are not mentioned here, but it suffices to say that for our purposes they will not play any significant role.

2.6.2 Semantics

Semantics is devoted to the study of meaning, as witnessed at the levels of words, phrases, sentences, and larger units of conversation and discourse. It focuses on the relation between signifiers, like words, phrases, signs, and symbols, and what they stand for and their denotation. Linguistic semantics is the study of meaning that is used for understanding human expression through the vehicle of language. The study of semantics is also closely linked to the subjects of representation, reference and denotation. The basic study of semantics favors the study of the meaning of signs, and the study of relations between different linguistic units and compounds listed below:

- Homonymy: one of a group of words that share the same spelling and pronunciation but have different meanings.
- Synonymy: a word with the same or similar meaning of another word.
- Antonymy: one of a pair of words with opposite meanings. Each word in the pair is the antithesis of the other.
- Hypernymy and hyponymy: a hyponym is a word or phrase whose semantic field is included within that of another word, its hypernym.
- Meronymy: denotes a constituent part of, or a member of something.
- Metonymy: a figure of speech in which a thing or concept is called not by its own name but rather by the name of something associated in meaning with that thing or concept.
- Holonymy: defines the relationship between a term denoting the whole and a term denoting a part of, or a member of, the whole.
- Paronyms: a word that is a derivative of another and has a related meaning.

Re-creating true semantic understanding has also been, with good reason, considered one of the key paths toward intelligence in machines [61, 62]. Unfortunately, this goal despite its importance, was not truly attainable given the constraints of the experimental design and setup conducted for this doctoral study.

2.7 Memory

Discussed in this section is memory in biological systems and the distinction between associative and working memory. Associative memory is simply the ability to formulate an association with an object or an action that can be used at a later state to recall one instance when presented with an appropriately paired instance or object. This is a very important basis for learning since it provides recall which can, in theory, be used to develop higher-level concepts. A working memory construct that could engender learning is proposed in a paper by Izhkevich [63]. The model is biologically based, as one would expect, and serves as a reference point with regards to behavioral properties, but not actual physical design. In Sections 2.7.1 and 2.7.2 more about the specific types of memory relevant to this work is presented in detail.

2.7.1 Associative Memory

Associative memory refers to a connection between conceptual entities as a result of similarity between those states or their proximity in space or time. Memory seems to operate as a chain of associations: concepts, words and ideas are interlinked, so that stimuli such as a favorite toy will call up the associated name [64, 65]. Understanding the relationships between different items is fundamental to episodic memory.

Classical conditioning is an example of associative memory driving the learning process. In the famous experiment, Pavlov paired the sound of a bell with food, and later the dog salivated to the bell alone, indicating that an association had been established between the bell and food [66, 67, 68].

For operant conditioning we observe behaviors increase in strength and/or frequency when they are coupled to a reward. We believe that this follows from an association between the behavior and a mental representation of the reward.

There is no specific “reward” in this work, just a pre-programmed attentiveness that will give us a greater probability of viewing interactions and events that will lead to linguistic concepts.

2.7.2 Semantic Memory

Semantic memory is the memory of meanings, understandings, and other concept-based knowledge, and is directly related to the conscious recollection of factual information and general knowledge about the known world and imagined world. Semantic and episodic memory together make up the category of declarative (or working) memory, which is one of the two major divisions in memory. Semantic memory helps us to give meaning to otherwise meaningless words and sentences. We can learn about new concepts by applying our knowledge learned from things in the past [69].

Semantic memory includes generalized knowledge that does not involve memory of a specific event, but more the assimilation of several memory events. That is, semantic memory contains information about what a paper is, whereas episodic memory might contain a specific memory of writing this paper. Semantic memory is also where we can hopeful begin to discern the early stages of more intelligent systems. No experiments for semantic memory were conducted during this study due to temporal requirements of this experimental design.

2.8 Neuromorphic Circuitry

The hardware aspects of this problem naturally led us to consider neuromorphic circuitry. Neuromorphic circuitry is the study and development of novel electronic and optical technology that is motivated by the biological organization and function of the brain and its neural structures. Brain in this context usually refers to the human variety, but many types of brains have been studied and modeled with circuitry. It is a relatively new field in electrical and computer engineering tracing its origins back to the end of the 1980s and Mead and his seminal book on the topic [70] and the paper that followed [71]. It is a highly interdisciplinary area of research requiring understanding and collaboration between electrical and computer engineering, psychology, neuroscience, neurobiology, mathematics and computer science to name a few. In addition to the desire for novel devices and architectures neuromorphic circuitry is being used as a tool in reverse engineering the human brain, a grand challenge for engineering according to the National Academy of Engineering. This is important for many reasons, but we will focus on the aspects that are perceived to be closest to the goals of electrical and computer engineering.

As engineers, we would be foolish to ignore the lessons of a billion years of evolution. – Carver Mead

Noise is generally seen as an enemy of the scientist and engineer. It clouds data, obscures results, makes our research lives unusually challenging. As a result, we have developed many tools and methods to combat noise. We utilize metrics such as the signal-to-noise ratio and bit error rate to quantify and eventually reduce noise. We make a concerted effort to operate our theories and our devices in regimes that are as devoid of it as possible. As we push further into nanotechnology, we enter a land where our battle against noise may very well be futile. Moore's law is truly doomed if we do not become accustomed to the realities of noise and change the conversation to fundamentally address alternatives that utilize this inherent and natural state of the world we live in. Neural systems, however, have long adapted the ability to operate in very noisy environments and can guide us to new possibilities in design. All of this is possible with what would be considered unreliable circuits by the standards of electrical and computer engineering.

Another highly coveted attribute of the brain is its ability to do all of this computation with very low power requirements. This peaks our interests in an age where more technology is gathered on a single device or system with the general expectation of reduced form factors and increased battery life.

Mixed signal architectures have also started to play a more important role in the design of these circuits. Many initial efforts were focused on analog signals mainly due to the brain's own analog architecture. But it has been shown that there is merit in the digital incarnations of neurons as well. Complementary architectures that hold true to the neuronal structure, i.e. very large fan-out/fan-in, inhibitory and excitory signals, are key components of neuromorphic circuitry. Developing computational models that work within this framework is equally important for the progression of the field.

2.8.1 The Silicon Brain

Neuromorphic circuitry is being pursued by several groups, Boahen at Stanford and Modha at IBM Almaden to name a few. Boahen's group has developed the Neurogrid chip [72] and architecture which allows them to simulate massive numbers of neurons in-silico [73]. The hope is that in the future the Neurogrid in Figure 2.3, will become a prototype for affordable supercomputing.

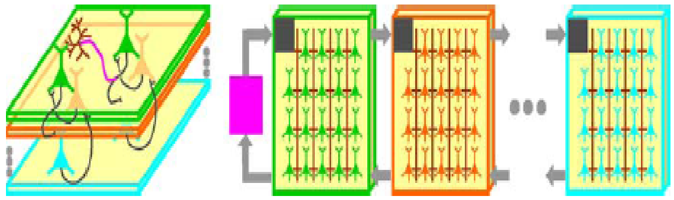
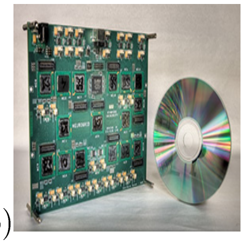
a)  b) 

Figure 2.3: (a) Neurogrid, a platform for cortical simulations. Cortical cell layers (left) are mapped onto chips (right) with arrays of silicon neurons. (b) Finished Neurogrid FPGA board. Image (a) is adapted [73], image (b) is from Boahen's Brains in Silicon website [74].

Modha is conducting research in a similar vein at IBM, but his group's focus is grander in scale. They are currently developing hardware and software to empower our current computing models with the parallel goal of achieving true cognitive computing. This is also developed with new types of supercomputing in mind. Modha's group has developed full-scale cortical simulations of a rat and cat scale simulations which are helping to develop our understanding of these systems [75]. This work is moving toward human neural constructs in silicon as well [76, 77].

2.9 FPGAs and Silicon Neurons

In a similar vein to Boahen's research, several groups have turned toward using Field Programmable Gate Arrays (FPGA) for the design and development of neurons in-silico. Spiking Neural Networks (SNN) have been a great source of inspiration for many of these systems. This bears a close resemblance to observed human brain processes and these groups hope to drive artificial activity by developing architecture with neural considerations from the start [78, 79, 80, 81, 82]. At this stage large-scale spiking neural networks have been achieved with this technology [83] and study continues in the search of alternative computing architectures.

2.9.1 FPGAs versus ASICs

Application Specific Integrated Circuits (ASIC) and Field Programmable Gate Arrays (FPGA) are both useful tools in the design of hardware-based neural networks. Each of these technologies have different advantages and drawbacks. For example, using FPGAs allows for quick updates, but is slower than using digital ASICs; using analog techniques implies real-time and often lower power operation, whereas using digital techniques implies programmability and offers higher and controllable precision [84]. ASICs truly shine in cases where a general-purpose computer architecture cannot fully emulate the parallelism that neural networks are known to have. Research shows that the FPGAs currently do not match the ASICs in performance. The design of FPGAs is more cost efficient when testing and manipulating circuit design is required in real time. When considering the utilization of a FPGAs, software will still play an important role for the computing systems you wish to emulate. FPGAs provide the ability for customization in a cost-effective manner when designing circuits in general and artificial neural networks in particular. FPGAs are ideal for the cycle of design and testing of architecture, which can then be used to convert to ASICs when performance issues are paramount. Regarding the experiments conducted during this research, there were several initial attempts to introduce and design neuromorphic components to augment the ability of the iCub android. These were unattainable within the prescribed timetable.

Chapter 3

PRELIMINARY INVESTIGATIONS

Better that one heart be broken a thousand times in the retelling, he has decided, if it means that a thousand other hearts need not be broken at all– Elie Wiesel

3.1 Early Investigations

The following earlier investigations are included because they outline the portions of the mindset that lead us to choosing this topic and the power in subtle technologies chosen for this work. This dissertation would not be complete without this information.

3.2 Self-Organizing Map Simulations

In this section, we considered self-organizing maps, in particular, their neighborhood functions and how variations in this component maybe helpful in our study. Self-organizing maps can be treated as part of a simple neural network and the hope was there may be other types of applications for this technique in our work, i.e. visual-spatial compression, etc. SOM has been considered in motor babbling experiments proposed by a colleague, Dr. Lydia Majure. To achieve this goal, we explored several variations to standard SOM models with an eye toward investigating some of the weaknesses inherent in the model as it stands, i.e. time to converge, accuracy, etc. In particular, the experiments focused on how differences in training methods, learning rates, and map topologies decrease quantization error and improve clustering. This example was motivated by [85].

We used data on the national economies of 46 countries and upward of ten different economic indicators of financial health were added for each county, such as inflation, unemployment, trading balance, etc. The data was aggregated from open economic data sources and kept to one year's worth of information from the year 2000. The data was normalized across each indicator, so as to prevent a few indicators from being over-valued

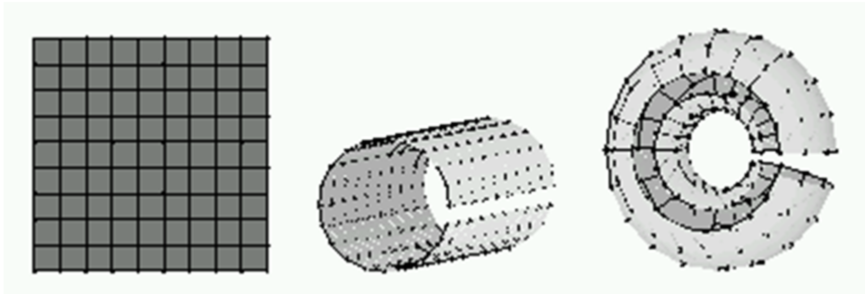


Figure 3.1: Flat (square), cylindrical and toroidal topologies.

or under-valued. The data was then processed using the basic Kohonen map algorithm in Section 2.3.2.1.

3.2.1 Topology

The standard grid layout is a hexagonal grid of nodes in a rectangular sheet. This topology presents some problems. The best neurons or nodes tend to gravitate toward the edges of the map, sometimes resulting in similar ones sticking on either end. To address this, we considered alternative topologies such as cylindrical and toroidal spaces, in Figure 3.1, that allowed node connections to wrap around the space and no longer reach a permanent edge.

3.2.2 Neighborhood Functions

The neighborhood function plays a significant role in the optimization process for SOMs. We tested the following standard functions, square wave, Gaussian, cut Gaussian and Epanevchicov, in Figure 3.2, to see quantitatively which nodes are changing on every temporal step of the algorithm.

3.2.3 Local and Global Optimization

A primary issue with standard SOMs is that only the nodes around the winning BMU are changed. This tends to create zones of local optimization, as seen in Figure 3.3. Early clusters tend to stay together and later samples have decaying effect on the maps predictions and updates. There is something positive to be said about this kind of robustness, but the ideal case would not be quite so resistant to new information. Preferably, there should also be an element of global optimization in this model. A simple

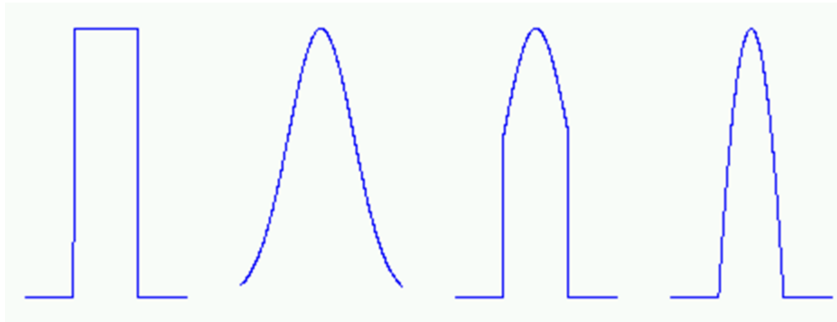


Figure 3.2: Neighborhood functions: Square wave, Gaussian, cut Gaussian and epanechicov (which is $max(0, 1 - x^2)$).

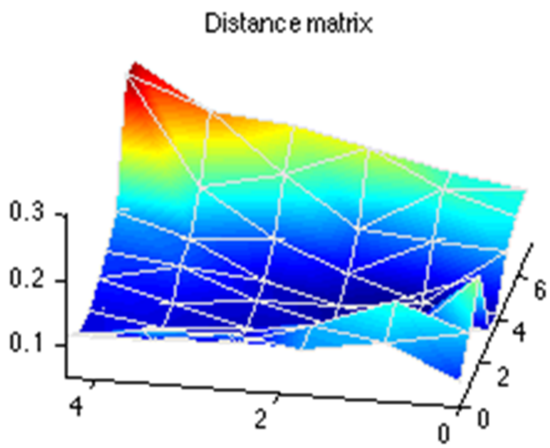


Figure 3.3: Optimization surface for the nodal distances. Red is least optimal and blue is most optimal on the z-axis, the x- and y-axes are spacial coordinates.

technique to vary this approach is to look at multiple winning nodes and allow them all to impact their neighborhoods. This will, of course, introduce an effect on the runner-up who now has less of an influence on its surrounding nodes. Subsequent testing showed a 4.79% decrease in quantization error. By itself, this statistic is not particularly significant, but it shows the potential benefits of global optimization.

3.2.4 Results

Cylindrical and toroidal grid topologies performed better than standard sheets. Updating weights of the network based on more global parameters rather than local neighborhoods prevented samples from locking in the network weights, although the training time is much

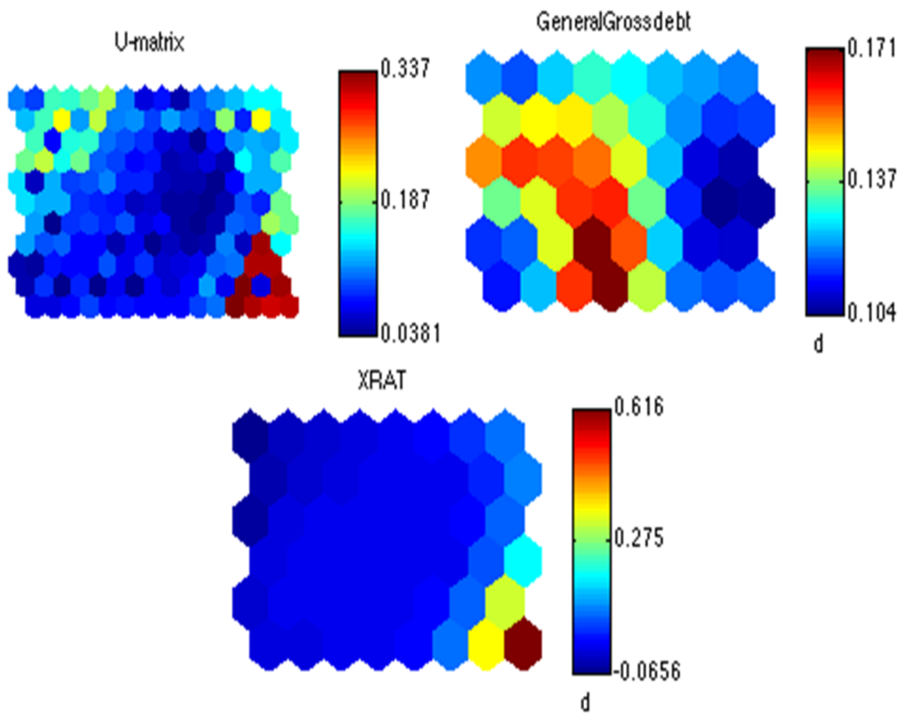


Figure 3.4: Country clustering by economic indicators. The color bars are the raw data values for each of the economic indicators.

higher. While not all combinations showed a significant improvement, many of them at least warrant further investigation. The first set of diagrams, Figure 3.4, show some of the clustering information derived from our first three economic indicators, the color bar denotes the distance metric from the accurate clustering of countries. It is important to keep in mind that the goal is to cluster unstructured data and in this case the SOM performs well as expected.

Figure 3.5 shows the last two economic indicators and the distribution of the countries based on the grid.

We can see that the method is able to conduct the self-organized selection as expected and our improvements do not prevent this from occurring. We can also see that we can overlay text association with the self-organized information even at this stage. These attributes will prove important as we proceed.

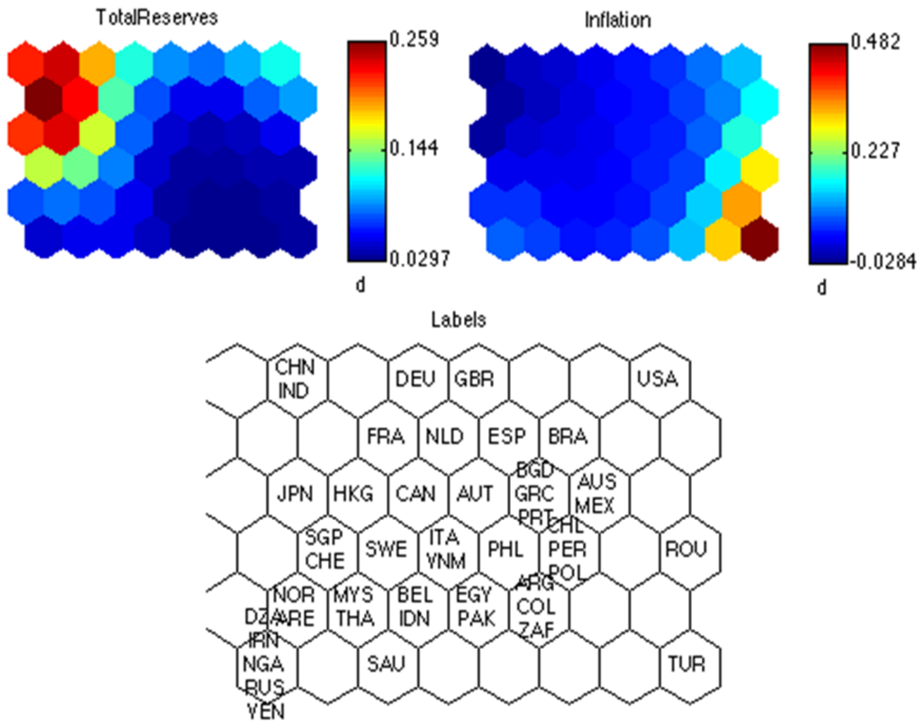


Figure 3.5: Country clustering by more economic indicators. The color bars are the raw data values for each of the economic indicators. The last image shows the external association that can be made to these clusters.

3.3 Simple in-Silico Neuron Design on FPGA

The motivation for this experiment was to consider hardware development in the process of implementing this framework, i.e. adding computation power to the android. As a result, we made attempts to be able to replicate a simplified neuronal model on an FPGA using VHDL logic. Once the artificial neuron is created, self-organizing algorithms (Kohonen maps) were implemented and checked for aberrant behavior. Testing a neuronal model with a self-organizing map on the FPGA allowed us to investigate some of the behavior of these algorithms in alternative circuitry configurations. We simultaneously questioned how simplified can we make a silicon neuron. Is a self-organizing map a natural topological representation of these artificial neurons? We were not sure of the answer, but the question was compelling enough to continue. We first utilized various resources related to self-organizing maps, neural engineering and FPGA design. We worked on creating one functioning silicon neuron and then wanted to scale up to as many as our Digilent Atlys FPGA could handle. We proceeded with a design from another group that worked on a similar problem.

ANNs, at this stage, are not as powerful or efficient as the biological networks that inspired their design, but both biological and silicon neurons have been shown to be more efficient with power and space than digital computers [86]. We understand that the brain is a massively parallel and efficient information processing system. Parallel processing can be utilized when using the FPGA design. We want to be able to gain better performance of neural networks with less power consumption. Clearly, this required a significant reduction in the overall functionality of the artificial neuron in comparison to a biologically inspired neuron for practical computing purposes. Various areas like pattern recognition, function approximation prediction and robotic control are just a few of the applications where ANNs are utilized with great success [87]. But, there have been questions regarding how closely the brain can be replicated in digital/analog systems. The observation that the brain operates on some of the analog principles involved in the physics of neural computation that could be fundamentally different from traditional digital computing, helped to spark the field of neuromorphic engineering [88].

Silicon neurons are also another component that can be utilized in the replication of an artificial neural network. The silicon neurons (SiNs) are considered to be a hybrid between analog/digital very large scale integrated circuits, whose components are representative of the electrophysiological behavior of biological neurons and conductance based models. Silicon neurons in theory will eventually allow us to emulate directly neuronal computation in hardware rather than simply simulated on a general purpose computer which may have

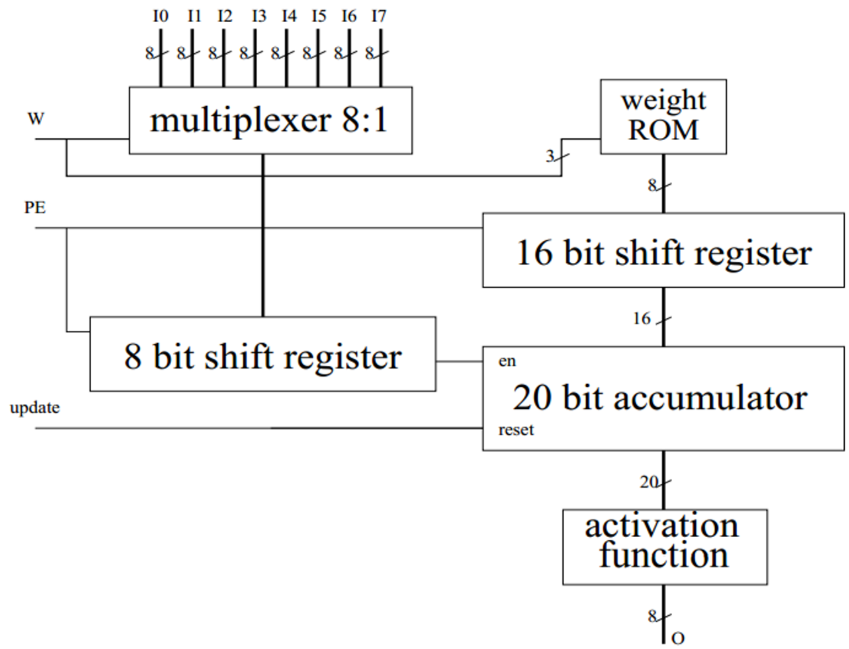


Figure 3.6: Neuronal schematic.

architectural constraints. Networks containing millions of neurons and ten billion connections, and complex models like spiking neurons with temporal information that require convolutions to be computed at each synapse, will possibly challenge even the fastest computers. Hence there is much interest in developing custom hardware for ANNs. SiNs will be most useful when large-scale dedicated neural computing is desired in real time and under stringent power and space/weight constraints, such as in neuroprosthetic, braincomputer interface, or embedded machine intelligence applications, as is the case in our laboratory.

3.3.1 Design Features of a Simple Neuron

The design features for our artificial neural network begins with the biological neuron. We understand graphically that need to replicate the change in voltage across the neuron in particular and that this could be done in many ways. Xilinx was used to program and debug in VHDL. Our simple neuron is shown in Figure 3.6.

We have our inputs entering the neuron and then summed together. When designing our neuron we utilized the block diagram provided in [89]. The block diagram contains a multiplexer, two shift registers, and accumulator weight rom and activation function. The shift registers, accumulator and multiplexer were simple to create. The activation function

Table 3.1: Estimate Sigmoid Function versus Actual Sigmoid Function

Input x	Output y_1	Output y_2
-2	0.119203	0.166667
-1.5	0.182426	0.2
-1	0.268941	0.25
-.5	0.377541	0.333333
0	0.5	0.5
.5	0.622459	0.666667
1	0.731059	0.75
1.5	0.817574	0.8

was modeled after the sigmoid function. It was difficult initially, because the sigmoid function has an exponential behavior in the denominator. This can make the sigmoid function is computationally expensive when done on the FPGA. We were able to discover a way to re-create the sigmoid function without the need to use a generated look up table.

$$y_1(x) = \frac{1}{1+e^{-x}} \tag{3.1}$$

The sigmoid function that will be used is shown below.

$$y_2(x) = \frac{1}{2}\left[\frac{x}{1+|x|}+1\right] \tag{3.2}$$

The components used to create this function were an adder, multiplier, absolute value and a divide-by-two entity. We wanted to verify that the function was a good estimate for the actual sigmoid function. In Table 3.1, we have a few input values that we compared against the equations above.

The percent error between the actual and the estimated sigmoid function increases when the input is a negative value compared to a positive one. The features of this network require the neurons ability to be interconnected and to continue to function in this state, this way the learning process will be operate as designed and the activation function used will convert the weighted input to a output activation. The design was for one neuron initially which was achievable and simpler to implement. This was a necessary step before we could begin modifying and manipulating the circuit in order to see how the neuron behaved before attempting to scale to an entire neural network. We also considered our space constraints on the FPGA in the design of this neuron in relation to scaling concerns. Off chip training was proposed by Salapura et al.[89] to reduce real-estate consumption iteratively during the design process. The activation function used, in that case, was the estimated version of the sigmoid function.

The self-organizing map requires that we determine the Best Matching Unit (BMU) and generally use a Euclidean distance in this process. Considering that the self-organizing map would be implemented on the FPGA as well, the Euclidean distance algorithm might be a heavier resource demand than we would like and is needed for our purposes since the Euclidean distance utilizes the square root function. To avoid having to use the square root function, which can be cumbersome, we used the Manhattan distance instead [90].

$$W(s+1) = W(s) + \theta(s)\alpha(s)(D(s) - W(s)) \tag{3.3}$$

where $D(s)$ = input vector
$W(s)$ = weight vector
$\alpha(s)$ = decreasing learning coefficient
$\theta(s)$ = neighborhood function $\sqrt{\sum_{i=0}^{n} x_i^2}$
x_i = member of data sample
n = number of dimensions in the sample vector
The manhattan distance is,

$$\sum_{i=1}^{n} |x_i - y_i| \tag{3.4}$$

where n = the number of variables and $x_{i,} y_i$ = the number of the ith variable in the x and y direction. The Manhattan distance is the sum of the differences in the x and y direction of two points.

The weights of the nodes are randomized at the initial operation. We then focus on the topological distance between the nodes or "neurons" effectively in our case. The neuron with the smallest distance to the input values in the topology is recognized as the winner. The change is then updated by a specified learning rule. This process continues until the entire neighborhood of neurons are approached with a similar methodology and then the weights are updated by the same learning rule that was used for the original weights. We also are aware that the strength of the learning should decay with distance. This decay is usually a Gaussian in nature. Figure 3.7 illustrates the flow through this process.

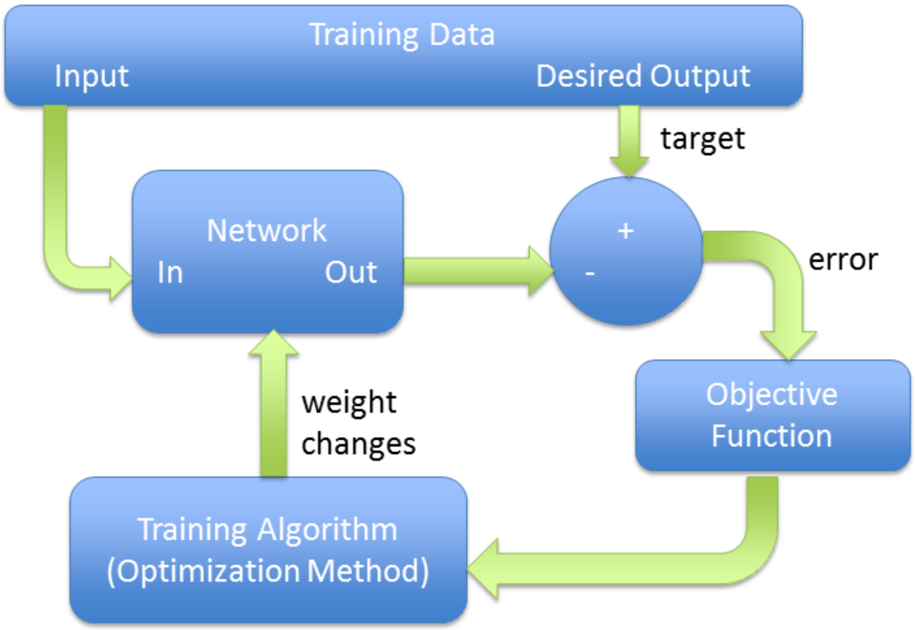


Figure 3.7: Neuronal flow chart.

3.3.2 Results

We realized that many of the simplifications needed for a neural network or neuron has made considerable progress over the years. The FPGA is a useful tool for the implementation of design and reconfiguration of artificial neuronal circuitry, but with limitations with regard to efficacy in transacting computational tasks. The type of neural implementations studied is the silicon neurons, application specific integrated circuits, and artificial neuron built using VHDL code for the FPGA. We have been able to create some of the entities required for our neural network. Further work needs to be done on the implementation side of this experiment. We have also started working on a learning algorithm that would be useful to use on the FPGA. The self-organizing map is a learning tool that will be integrated in the visual-spatial component of this framework, as expected. Analyses of the error back-propagation algorithm is another avenue to consider. As far as extending this functionality to the android, the case is not considered compelling at this stage. Future comparisons will be helpful in the decision process for the appropriate learning algorithms to use on our artificial neural network implemented on a FPGA. Unfortunately, due to budget constraints, we were unable to expand this work and dive deeper into some of the other questions we had regarding a successful integration with the android.

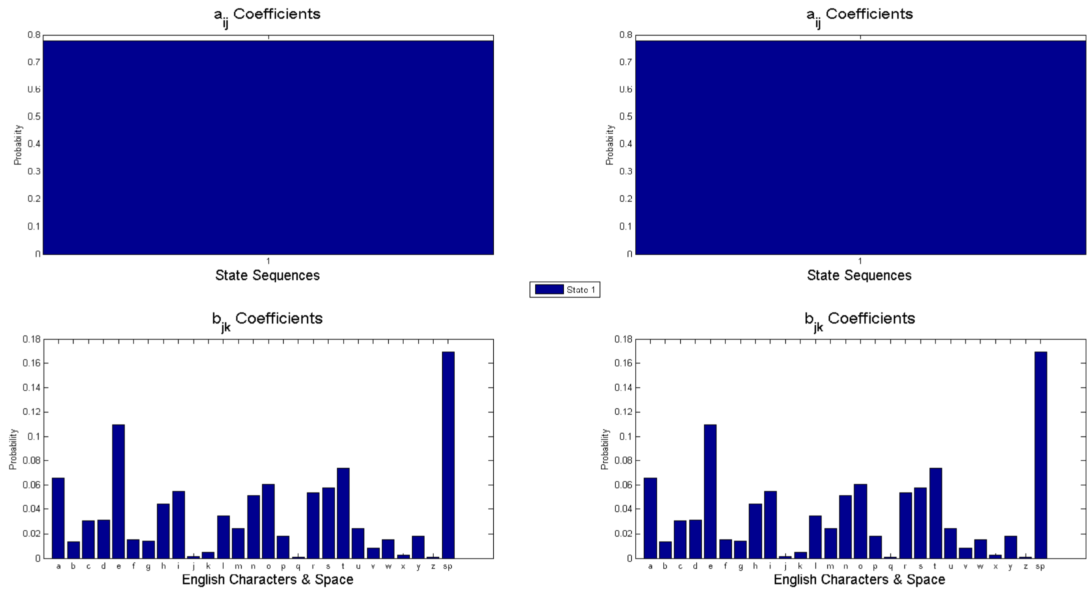


Figure 3.8: One state (trivial case).

3.4 Identifying Natural Language Structure via Hidden Markov Models

This research was conducted in parallel with Professor Levinson's Mathematical Methods of Language course at The University of Illinois at Urbana-Champaign. This was my initial research into the hidden Markov model with various states in an attempt to detect simple patterns in unstructured text. Collectively these are called the Cave-Newirth experiments which were conducted over a range of one to twelve states.

3.4.1 Results

We have displayed the results of the a_{ij} and b_{jk} coefficients which are also known as the state transition matrix and probability observation matrix, respectively. We started with the single case state, Figure 3.8, in order to make sure that the code was compliant with a basic and known result. We display this result for continuity.

The more interesting work begins in the higher-order states, Figures 3.9 - 3.11. As we increase the number of states thus allowing for more transitions, we see that the percentages for the alphabet increase as expected.

We still do not have enough detail from this graph to understand the underlying structure of the input we presented the system with. Our percentages are abysmal for any

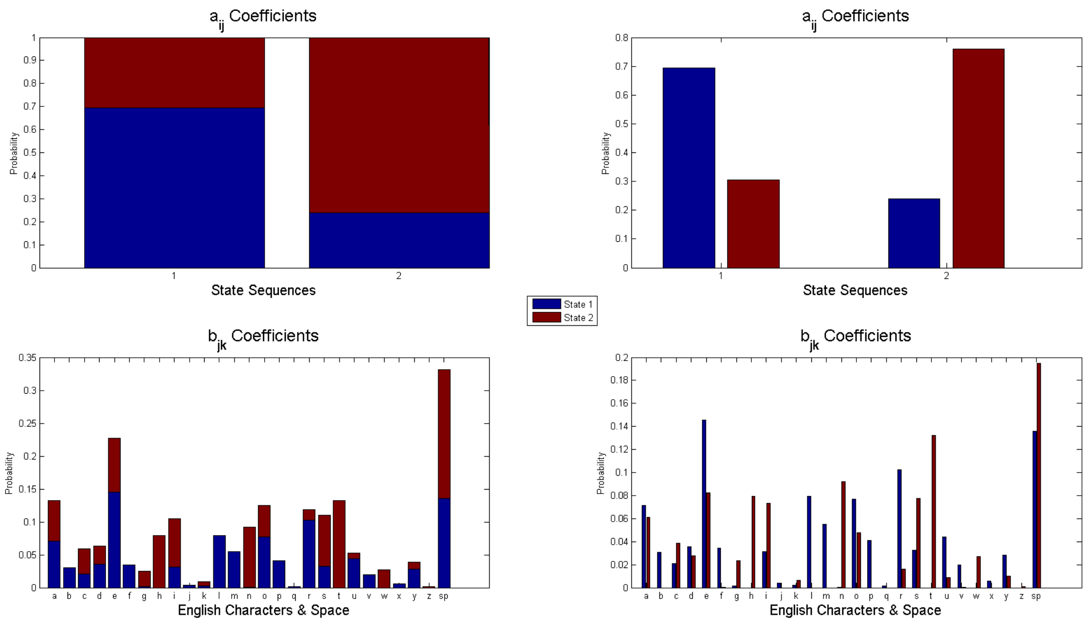


Figure 3.9: Two states.

individual letter in our lexicon. So we soldier on to higher numbers of states.

In a three state system, we make our first significant jump in the percentages reported in the states. Spaces are showing up a little less than chance in state three.

By the time we get to five states in Figures 3.12 we notice that the blank spaces are identified about ninety percent of the time. As we progress toward eight states we find that the first of the vowels, E, is identified one hundred percent of the time in the fourth state. The other letters that have greater than chance probabilities are H, N, R, S, and T. The vowels A, I and O are also steadily increasing in probability. Interestingly, the vowels U and Y are not making a strong showing yet. Another point worth mentioning is that the percentages for the letter H are above chance by the five-states system and actually dropped in the six-state system in Figure 3.13 by roughly twenty percent only to make gains in the seven- and eight-state system in Figures 3.14 and 3.15.

The other greater concern is the total time to completion for each system in Figure 3.20. We see a significant jump in total time from seven- to eight-state systems, Figures 3.14 and 3.15, and we can already see that the simulations are time consuming in the previous stages. Unfortunately, we were unable to log all of the data for up to and including the twelve-state experiment, Figures 3.16-3.19, but we are certain that given the trend and the increased requirements for computation, this would result in an increase in overall time. We are not sure about an approach toward ameliorating this issue, parallelization has been mentioned and might help to overcome this hurdle.

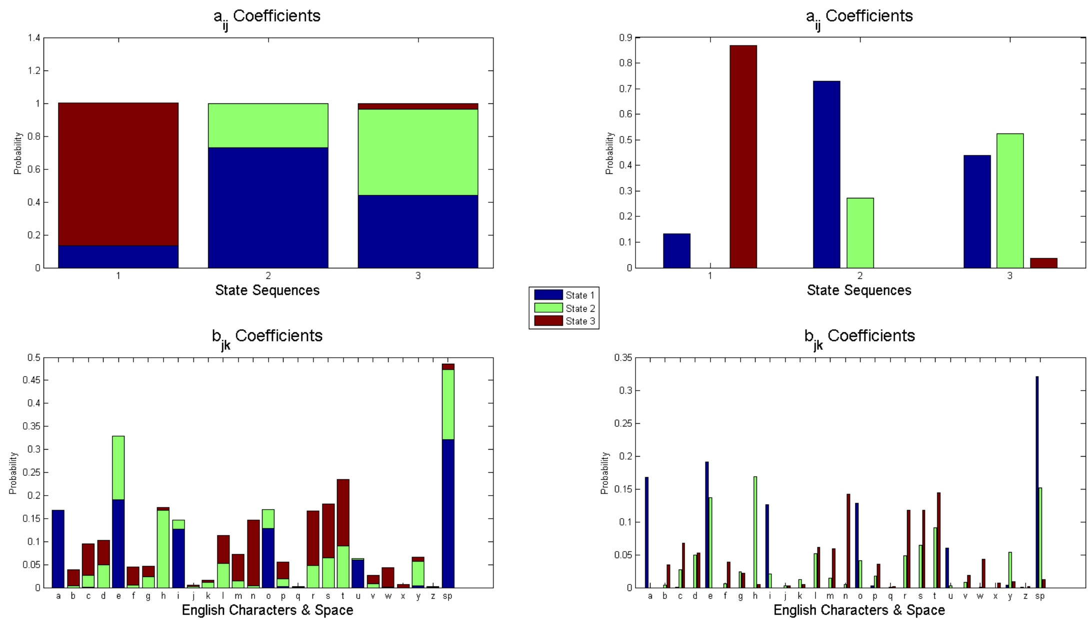


Figure 3.10: Three states.

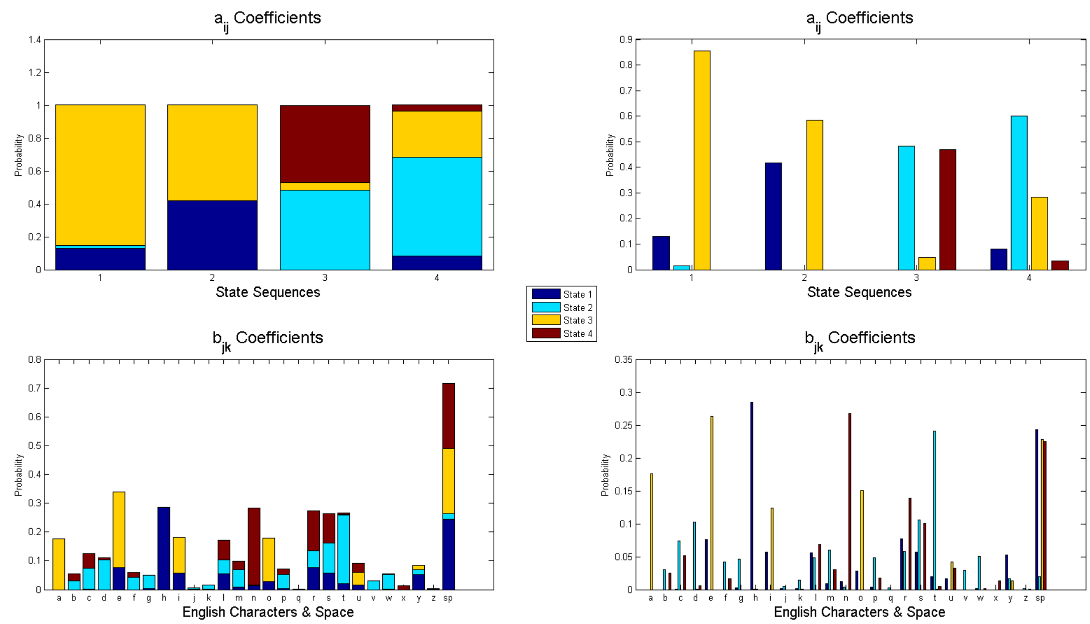


Figure 3.11: Four states.

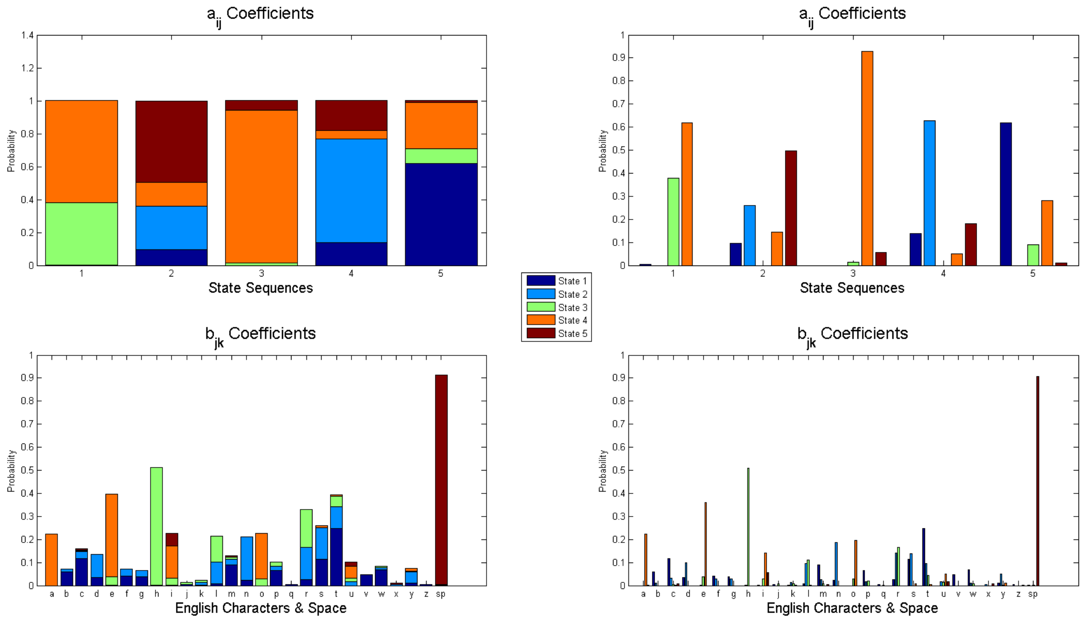


Figure 3.12: Five states.

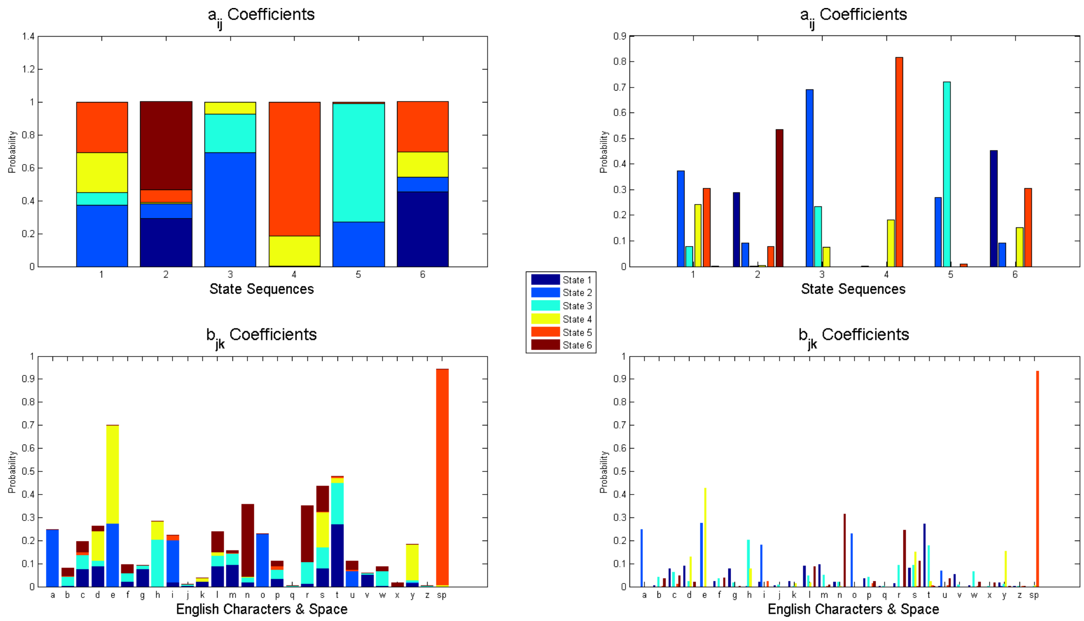


Figure 3.13: Six states.

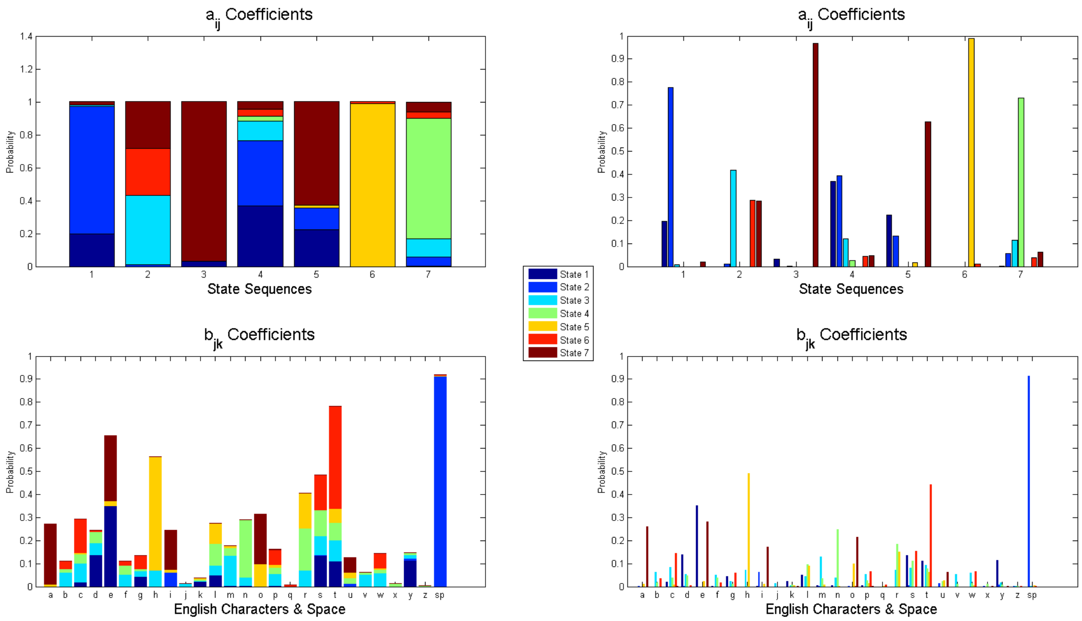


Figure 3.14: Seven states.

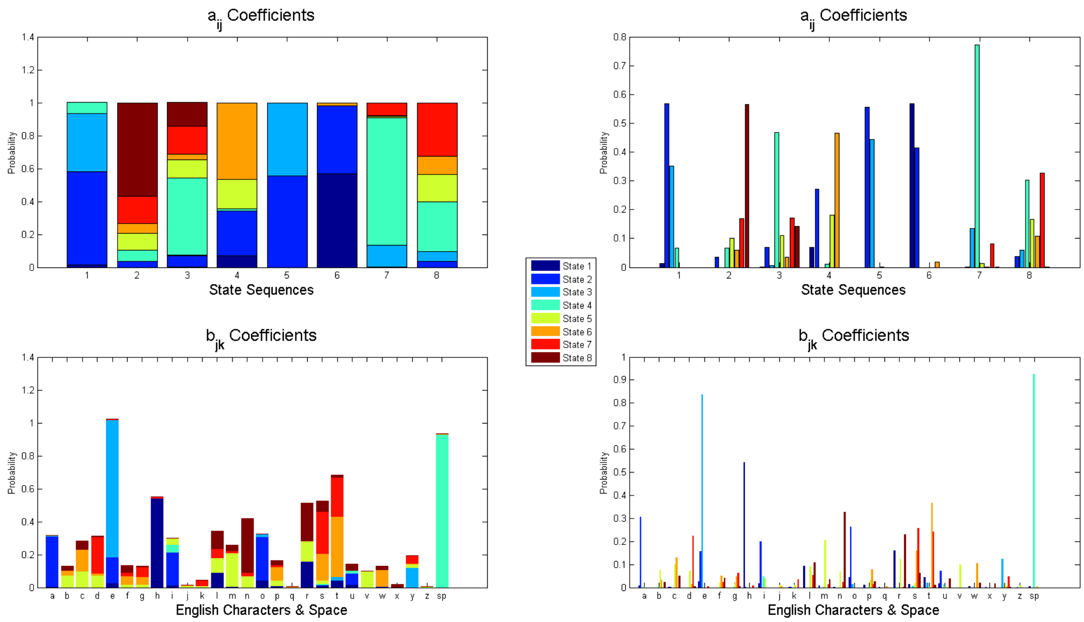


Figure 3.15: Eight states.

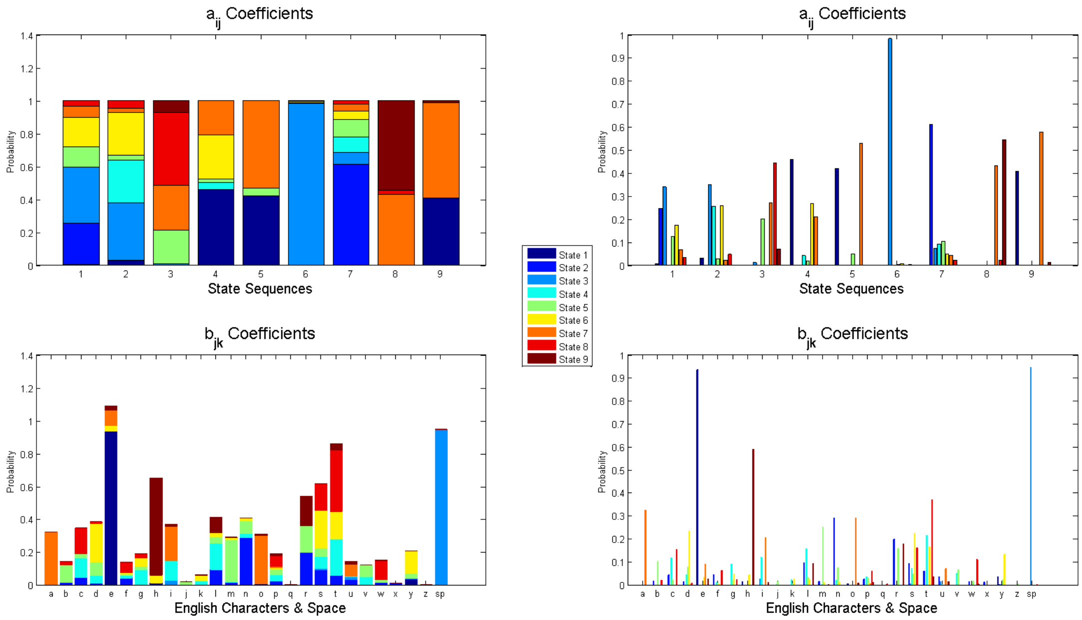


Figure 3.16: Nine states.

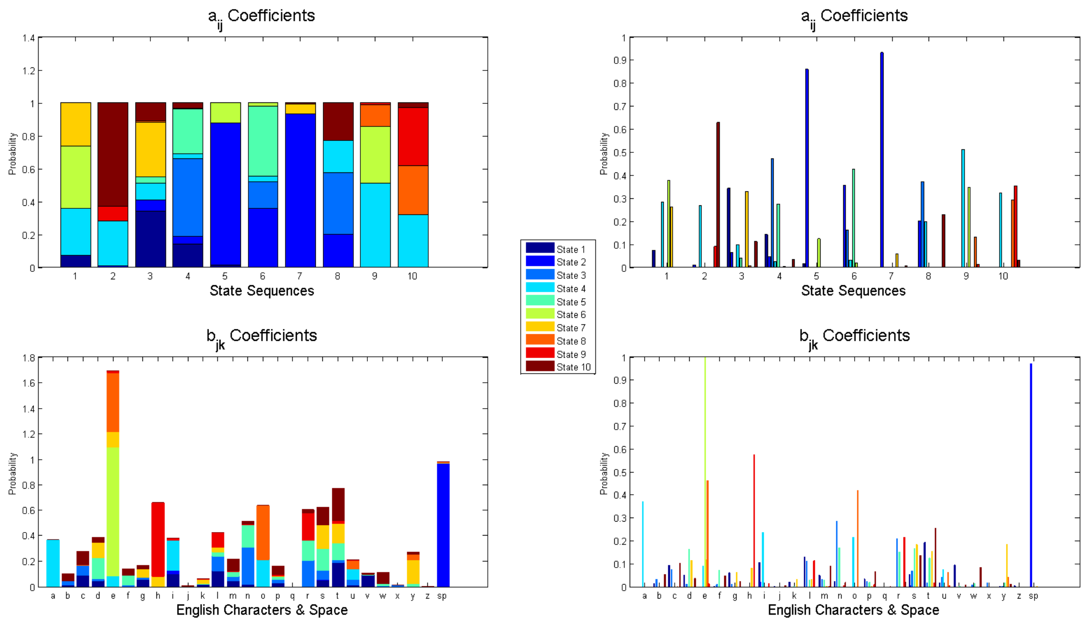


Figure 3.17: Ten states.

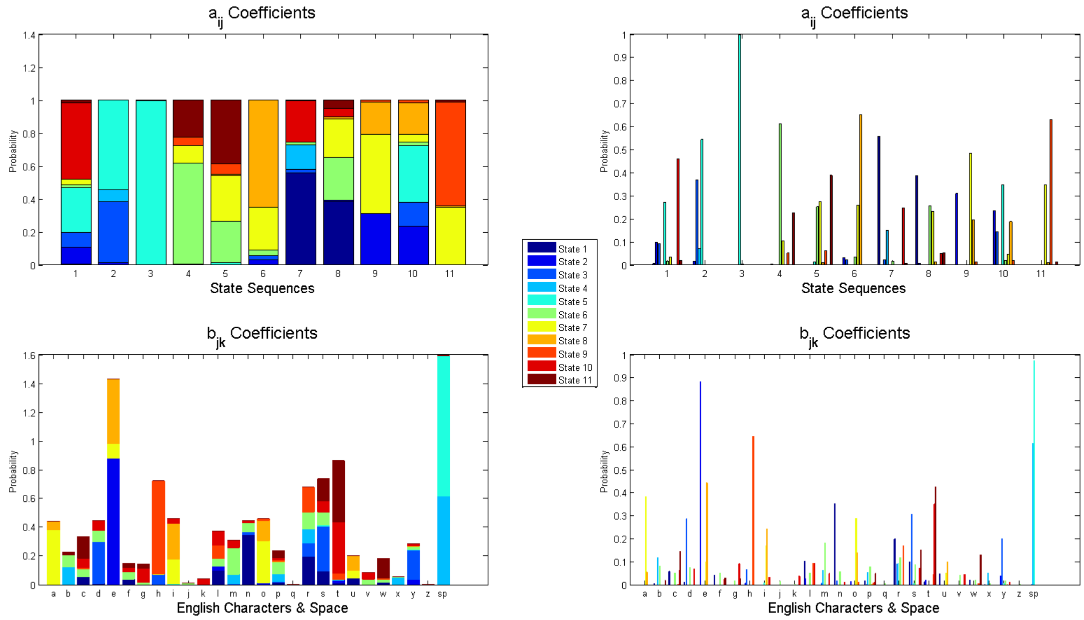


Figure 3.18: Eleven states.

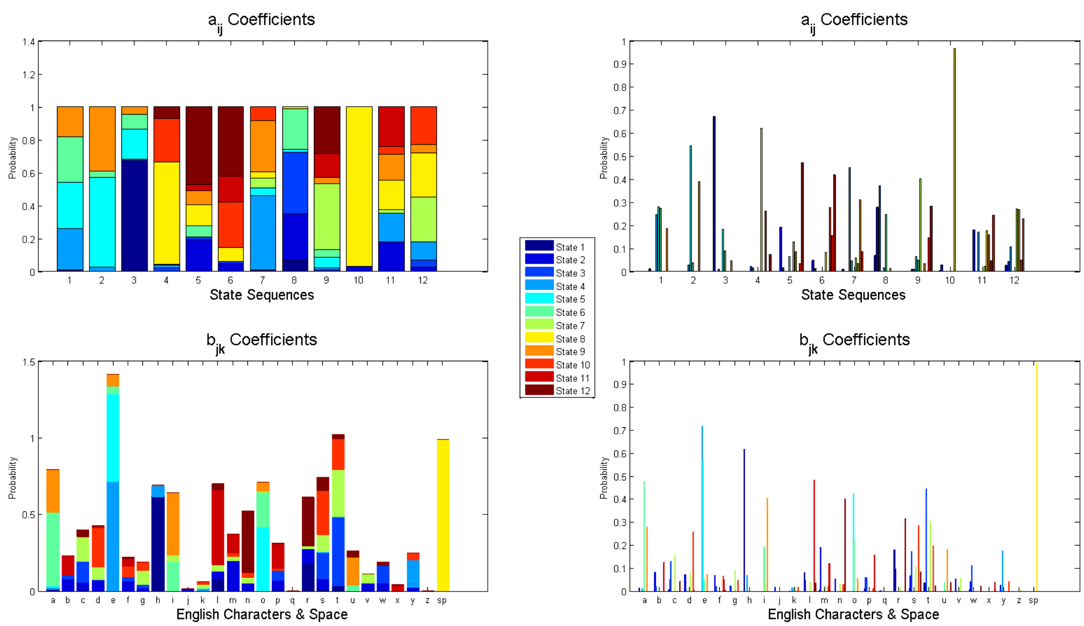


Figure 3.19: Twelve states.

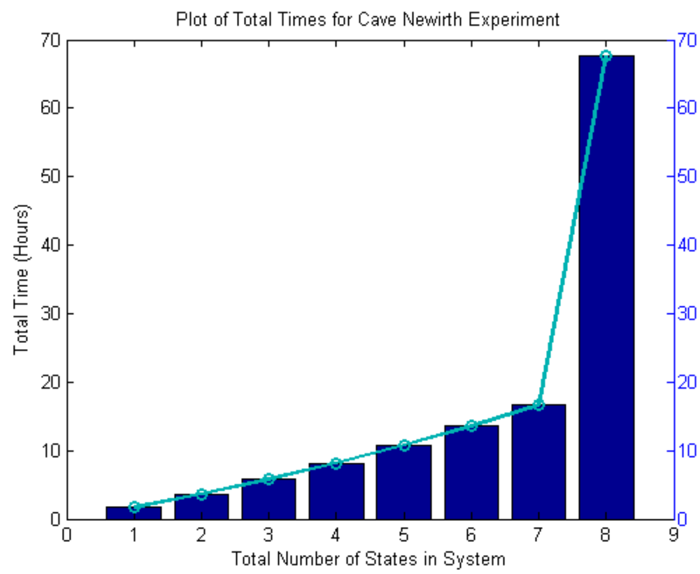


Figure 3.20: Total time for experimental runs.

We now talk about the convergence of the log likelihood function shown in Figure 3.21. As we expect, the system with the largest number of states converges is the quickest. What is interesting to note is that this case also has a far less pronounced wobble or plateau before it moves toward final convergence.

When all is said and done, the Cave Newirth experiment was intriguing to say the least. The hidden Markov model at work is very impressive at determining sequential structure where it is not immediately apparent. As we look through the results and start to see the distinction between consonants and vowels appear, we are taken aback knowing that this code has no prior knowledge of this fact. The next step is to apply this understanding to spoken language.

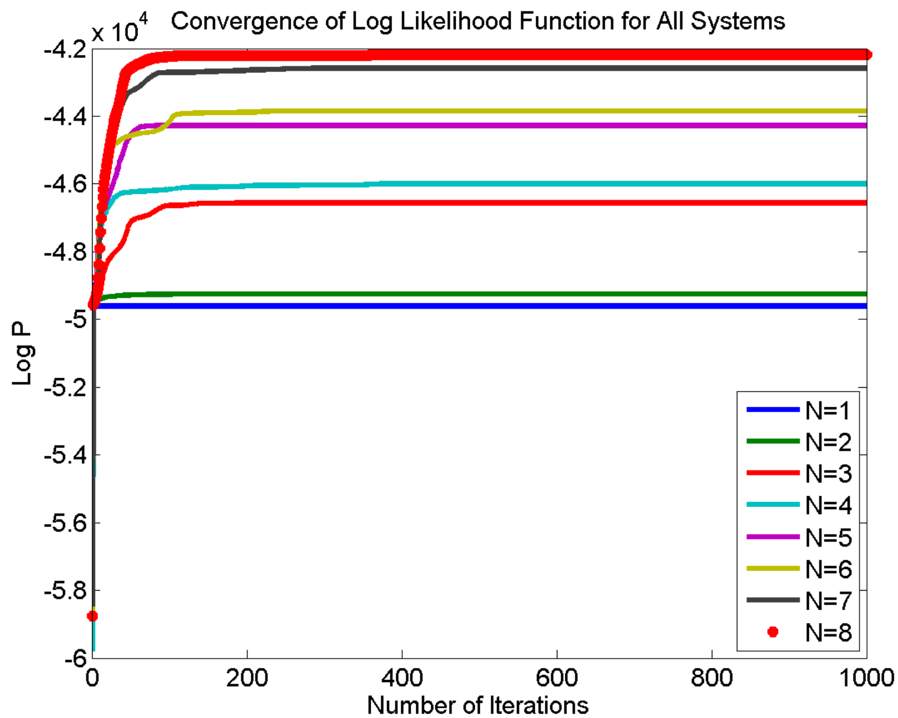


Figure 3.21: Log likelihood function for all systems.

Chapter 4

SYNTHETIC INSTINCTS AND CONCEPTUALIZATION

I hear and I forget. I see and I remember. I do and I understand. – Confucius

In this chapter the research methods used are discussed. In computational intelligence we have observed examples of unexpected and interesting behavior from very simple rules and drives implemented in simple systems i.e. swarm intelligence, etc. Usually the rules are extremely simple, stay a certain distance from their nearest neighbors, do not collide with their neighbor and so forth. We believe that simple rules are an excellent basis for greater complexity to emerge in more complicated systems, but we have to get the scale and nature of the rules to align with the appropriate environmental pressures needed to force a system to evolve or perish. Drawing from developmental psychology and observation, we applied a very small set (one or two) of synthetic instincts for the iCub robot that should support simple natural language acquisition. Swarm intelligence allows us to takes advantage of the knowledge of multiple external agents for information gathering and synthesis.

4.1 Synthetic Instincts for the iCub

We start with the synthetic instincts for the iCub robot. These instincts are simple autonomous action(s) from the android and provide some direction for some more complex supervised and unsupervised activities of learning that were implemented. More importantly we believe that providing a drive for the android to engage in events, real or imagined, is one of many steps toward observable intelligent behavior at any level.

Our android, Bert, is limited in its sensory capabilities. As a result, we have developed synthetic instincts, shown in Figure 4.1, that follow two criteria:

1. All instincts and drives are focused on gathering sensory information from the auditory channel, the visual channel or some combination of the two. For example, if Bert is confronted with a loud noise he should turn his head and focus on the largest moving blob.

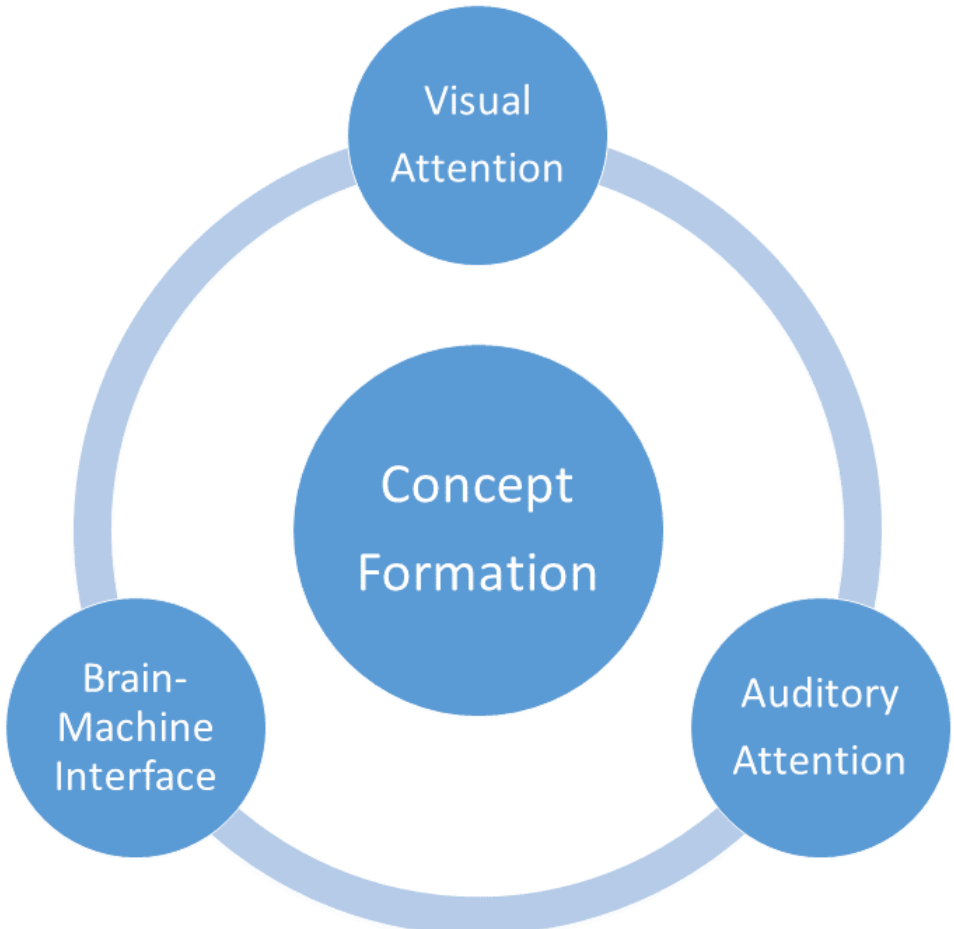


Figure 4.1: iCub synthetic instinct and sensory framework.

2. All instincts and drives are directly integrated in the process of acquiring language and/or communication capabilities, i.e. Bert may nod or smile after facial recognition in order to engage a speaker and attain visual or verbal cues for the accuracy of the information presented to him.

Following these criteria the synthetic instincts in the iCub are developed and utilized:

- Visual attention to all objects in frontal view like a newborn
- Auditory attention to speech-like sounds
- Babbling of words

4.1.1 Visual and Auditory Interactions

The visual instinct simply focus on what was presented to it at any given time. Implementing any extra tracking could have lead to conflicting data collection, so as a result we did not integrate that into this instinct. The auditory instinct will be trained to focus on sound, specifically on spoken word-like sounds. Visual and auditory inputs are then used to develop "concepts". These concepts are compared against previous interactions that the android experienced during the experimental trials.

4.1.2 Interrogative Drives

This set of drives is specifically designed to keep the robot engaged with the human interaction and to allow it to be an agent in the gathering of information. Keeping with the theme of simplicity, one instinct associated is to gesture in the western world for more information, or to make an auditory "hmmm" sound for more information. It was not necessary to fully implement these drives sinced the robot was engaged without the aid of these instinctual behaviors.

4.2 Conceptualization

Conceptualization is this study refers to the clustering and encapsulation of information for memory or linguistic interaction. Self-organizing (Kohonen) maps have been useful in the study of unsupervised clustering operations and the like. In this stage, the information gathered via the "swarm" sense is to be attributed to categories by feeding the input to the map and observing the natural clustering that occurs and then assigning the words uttered to the android during this time window to this new clustered data set that we will call a concept. Once a concept is "packaged", it can be made available to the linguistic apparatus in the android.

4.2.1 Simple Language Acquisition

The android, Bert, will need a mechanism for parsing and categorizing the auditory and visual information that he receives. The Hidden Markov Model (HMM) has been identified as one such mechanism from preliminary work. It consisted primarily of direct interaction with a robot, training it with word-object pairs. Providing cued video and audio input was

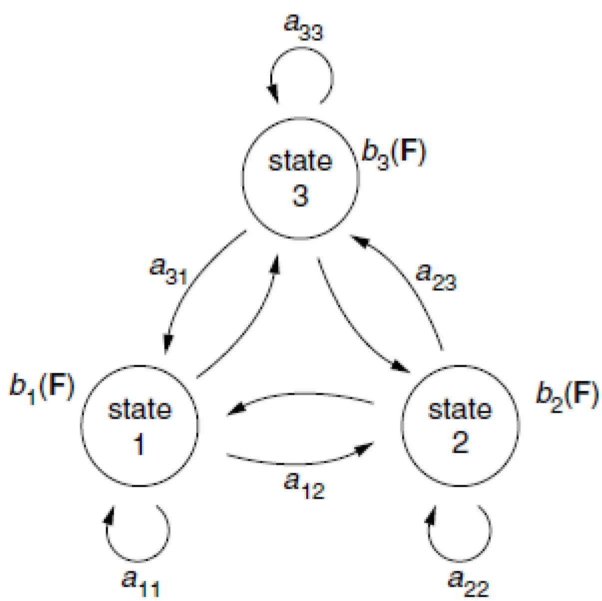


Figure 4.2: Three-state HMM. Adapted from *Mathematical Methods for Speech Technology.* [8]

considered as a secondary method of choice for developing word-object associations in the android system.

4.2.2 Human-Android Linguistic Interaction

At this phase in the research, the development and optimization of self-selecting algorithms for classifying linguistic structure are considered. The HMMs, specifically the discrete observation variant in this case, are well-known tools for this type of development [91].

The HMM, diagrammed in Figure 4.2, is a doubly stochastic process which is comprised of a state transition matrix, a_{ij} and probability observation matrix, b_{jk}. This machinery allows for the emergence of structure in previously unstructured input, text or speech in this case. The actual formulation is assembled in the forward and backward probabilities shown in the following equations:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}); \qquad 1 \leq t \leq T-1$$

and

$$\beta_t(i) = \sum_{j=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j); \qquad T-1 \geq t \geq 1$$

where O_t is an observation sequence. We now can write the following for the probability of an observation sequence conditioned on a model M.

$$P = Prob(O|M) = \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j) \; where \; 1 \leq t \leq T-1$$

This is collectively known as the Baum algorithm [8].

Chapter 5

SPEECH RECOGNITION

Every time we fire a phonetician/linguist, the performance of our system goes up. – Fred Jelinek

In order to learn language, a system needs to be able to distinguish speech signals from the many different types of auditory inputs it receives, often simultaneously, from background noise. CMU Sphinx, from Carnegie Mellon University, is an open source software package that has gained some fame for speech recognition [92]. The Sphinx-4 system is a flexible hidden Markov model-based speech recognition system. Its components can be configured at runtime along the spectrum of semi-to-fully-continuous operation. This research was solely concerned with continuous operation as we attempt to keep as much of the interaction in real-time as possible. The package also includes a suite of speech recognizers, pocketsphinx 4, and an acoustic model trainer, SphinxTrain, that were not needed for this research. A diagram of the system architecture can be seen in Figure 5.1.

5.1 CMU Sphinx Recognition Structure

The CMU Sphinx recognition system works in several sequential stages which are listed in Table 5.1. It has the following seven stages:

Table 5.1: CMU Sphinx speech recognition stages.

Segmentation, classification, and clustering
Initial-pass recognition
Initial-pass best-path search
Acoustic adaptation
Second-pass recognition
Second-pass best-path search
N-best rescoring

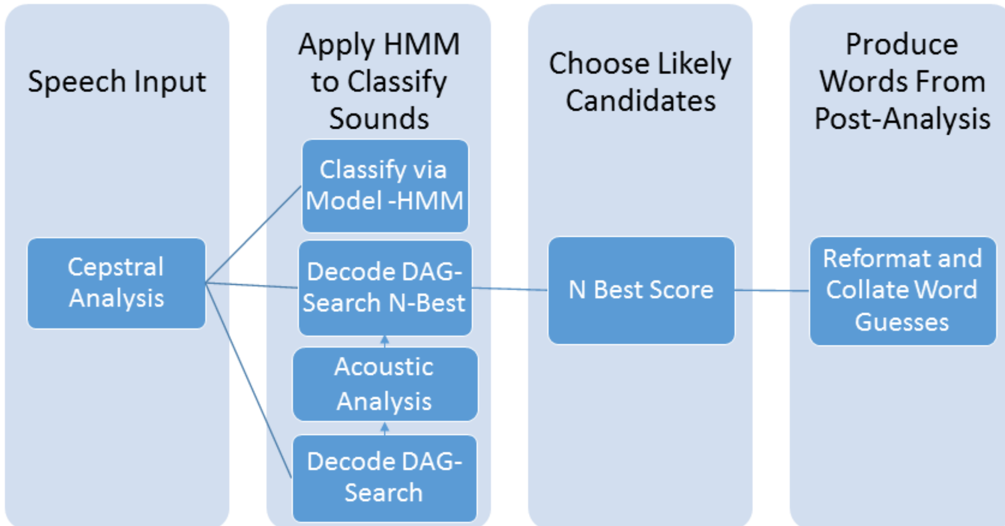


Figure 5.1: CMU Sphinx architecture.

In the first stage, the audio streams will be lumped into smaller segments. The segmentation will land such that they coincide with acoustic boundaries, e.g. silence. During the initial-pass recognition, the recognition is first done with a straightforward continuous-density Viterbi beam search. This will produce a word lattice for each sub-segment of the auditory input. Then we have the initial-pass best-path search. We search the lattices for the global best path according to the trigram grammar we have specified. We apply some acoustic adaptation via the HMM which is then adapted using Maximum Likelihood Linear Regression (MLLR). This adaptation is performed in one shot with a single regression matrix. We give the data a second pass for recognition here, each sub-segment is then decoded like the first time, using the acoustic models adapted in the previous step. A lattice is produced for each subsegment as expected. The second-pass best-path search involves searching the lattice for the global best path and an addition N-best search over the lattice is also done. Finally, an N-best rescoring requires the N-best lists generated using the supplemented vocabulary to be processed to convert the phrases and acronyms into their constituent words and letters, which we hope are accurate, but even errors at this stage could be useful for our goals.

5.2 The iCub Simulator

The iCub suite comes with a fully-functional simulator that can take true real-world inputs and act on them in a virtual world. This world can have objects with real physics thanks to the ODE package integrated in the release. The research and the experiments are conducted in the simulated environment, see Figure 5.2, without any loss of realism because the video and audio feed are both from the real world. The actual android in our laboratory currently runs on a version of the software that is several generations behind the version in which the code for this research was written. In addition to the real risk of damage involved in the operation of the android and some other mitigating circumstances involving the microphones on the android, it was decided that since there is no variation in the experiment from real to virtual environment that the entire experiment would be run in simulation.

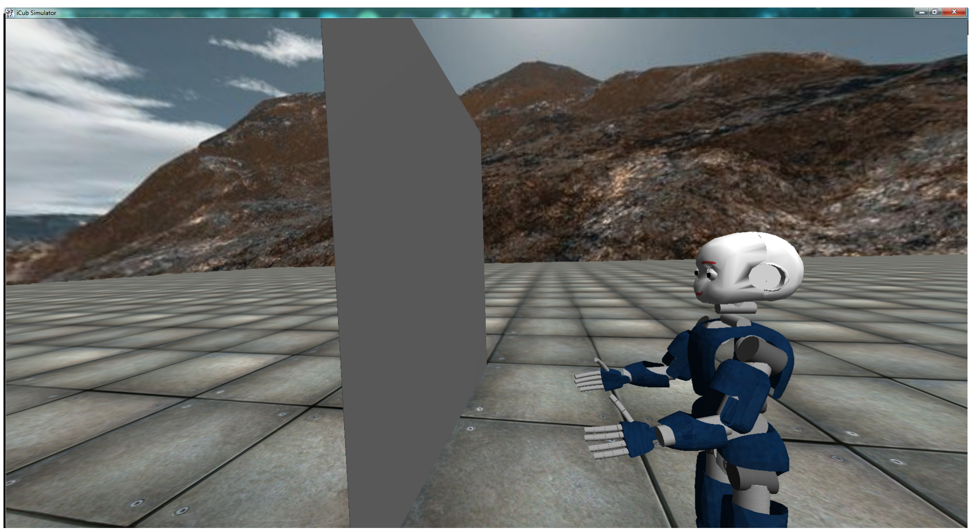


Figure 5.2: iCub simulator with a screen that allows the iCub to see video and interact with the real world.

Chapter 6

SENSORY FUSION AS EMBODIED COGNITION

We categorize as we do because we have the brains and bodies we have and because we interact in the world as we do. – George Lakoff

Our iCub Bert is a unique android built in the image of a 4-year-old child. Unfortunately, he is not equipped with a wide range of sensory input modes at this time. It is possible to collect visual and auditory information with accuracy. The most important aspect of the idea of embodiment in our research is the confluence of sensory inputs and the emergence of states that can be fostered by having access to this data. Language also has its very roots based as a tool for interacting with these senses [93]. Section 6.1 discusses the implemention of sensory capabilities beyond the visual and auditory in the android.

6.1 Utilizing Extra-Sensory Modalities

Keeping in mind that this is a humanoid system, senses that are above and beyond the norm might be helpful in developing semantic memory. The Microsoft Kinect system provides an infrared and depth sensor in addition to the ability to program algorithms to extract data directly from the system. Figure 6.1 shows the model used in the experiments with the android.

6.2 Brain-Machine Interfaces

The electroencephalograph (EEG) has in many ways become the de facto brain machine interface. Given the techniques temporal accuracy and reliability coupled with its generally non-invasive nature it is a perfect fit for experimental and commercial uses. Research in this area has been active for decades, making this a robust means of data collection with several companies developing commercial models for everyday use. Used in this research is the Emotiv EPOC+ EEG, seen in Figures 6.2 and 6.9, for data collection from 14 channels,

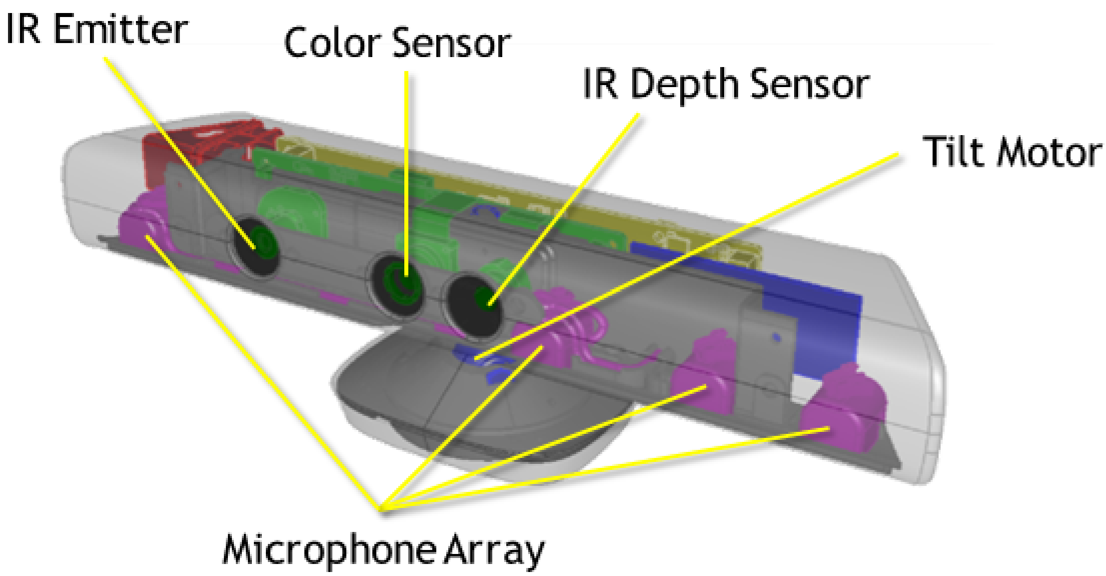


Figure 6.1: Microsoft Kinect.

AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4 across the entire head. It has a sampling rate of 128 SPS or 256 SPS running on a 2048 Hz internal clock. The Emotiv also comes equipped with software that also tracks mental focus, engagement, interest, excitement, affinity, relaxation and stress level of the subject.

6.2.1 Emotional States

In Section 1.2 we see that emotional states play a significant role in the development and acquisition of language. Using the Emotiv EPOC+ we track the subject's emotional and attentive state in real-time while the utterances to the android were made. This information allows the system to develop a fuller picture of the word meaning and association.

6.2.1.1 Affectiv Suite

Collection of emotional responses that are tracked via the EEG. These responses are the following: Frustration, Engagement/Boredom, Meditation, Excitement and Valence.

Figure 6.2: Emotiv EPOC+ EEG.

6.2.1.2 Effectiv Suite

In addition to emotional data, the EPOC+ can also be used to guess facial expressions in real time. This provides another window for embodiment data to be fused with language interaction and acquisition.

6.3 Swarm Agents

The swarm agents were programmed with a very simple set of directives. The agents are given a word and based on the emotional content they will move through the android's memories and choose other words of similar emotion content as determined by the EEG data that is saved for all of the linguistic interactions with the system. The number of agents available at any particular time will vary to reflect the variability of responses in any given situation.

6.4 Memories

The previous sections help to lay the groundwork for associative memories. CMU Sphinx is used to help provide Bert with the feedback required for word recognition. Associative memory is a useful component for conceptual construction and language acquisition.

Figure 6.3: The experimenter in the EEG EPOC+ apparatus.

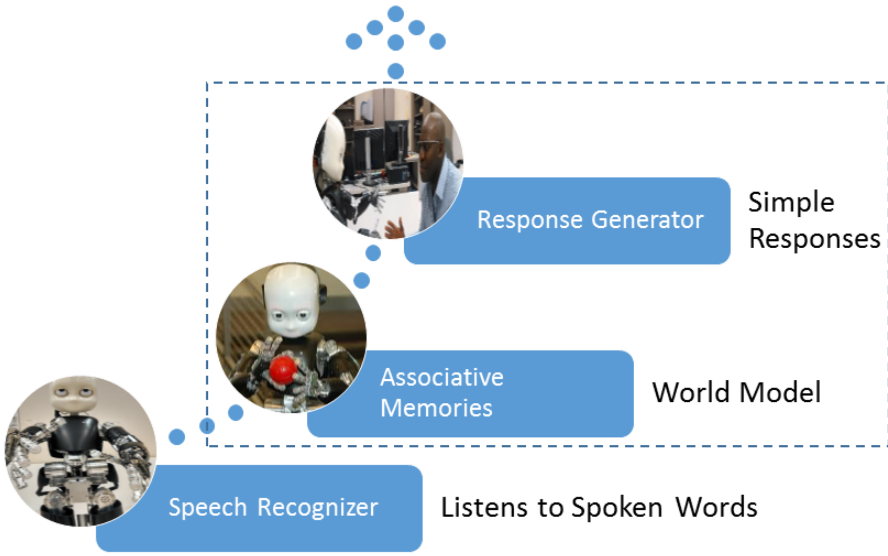


Figure 6.4: The data and response flow. The box indicates the contribution in this sequence for this research.

The work focused on the development of the components within the dashed box shown in Figure 6.4.

Chapter 7

NEURAL NETWORKS

Artificial Intelligence: The art of making computers that behave like the ones in the movies. – Bill Bulko

7.1 Neural Network Phase

At this phase, the sensory data gathered is used to reach the goal of this research which is to acquire simple language skills, e.g. babbling after interactions with another agent.

7.1.1 Self-Organizing Maps

The linguistic information was gathered during casual talking and the android was coupled to emotional data, simultaneously, via the EPOC+ EEG. At this stage, this data was clustered via a Kohonen map or so-called Self-Organizing Map (SOMs). This provided emotional clustering for the verbal data. The visual component of this experiment was accessible via associative memories stored as image files.

7.1.2 Echo State Network - Reservoir Computing

We have allowed the interactions to cluster and we let the android improve its babbling by attempting to learn from these and future interactions. A reservoir computer is a neural network that allows for the implementation of a model that does not require the user to provide the specifics of the learning task. For more details, please refer to Section 2.3.2.2. The goal for our android was to proceed via trial and error which is very similar to the way language is learned by human beings. As a result of this method, the babbling we observed for this experiment would require a significantly long time commitment on the part of the experimenter if obtaining a full fledged linguistic response was attempted. But at this stage, it is possible to verify emotional alignment and contextual relevance for the babble produced over time versus stochastic babbling production.

7.2 Speech Production

For the speech production portion of the work another well-established open-source software package called Festival (or FestVox) was used. This is the recommended package for the iCub and it works fairly well.

7.2.1 Linguistic Babbling

As with all creatures or systems that are learning there will be many imperfect attempts made at communication. These types of responses were observed and fully expected during our interaction with the android.

Chapter 8

INSTINCT-CLUSTER-SWARM-MEMORY

The best way to predict the future is to invent it. – Alan Kay

8.1 Instinct-Cluster-Swarm-Memory Formalism

The synthetic instincts are based on the android's most accessible modalities, vision and audition. Unifying these sensory streams required a reliable metric, in this case timing was sufficient, and once we completed this the integration of the swarm agents was treated as another "extra-sensory" modality which provided adaptation for the model. The formation of the "concepts" which are used in the linguistic attempts during the experiments. Figure 8.1 displays the research paradigm followed in this work.

8.1.1 Instinct

Synthetic instincts are treated in an axiomatic fashion during this work. There was no a priori expectation of proving why they were selected, but one can argue that their presence motivates the desired behavior.

Let us consider the following:

We are given a set of input streams for the iCub, $X, Y, Z \in \Omega$ in $\mathbb{R}^n$, relating to visual input, auditory input and data stored in a repository respectively. They are discrete in nature,

$$X = \{x_1, x_2, \ldots, x_n\} \tag{8.1}$$

$$Y = \{y_1, y_2, \ldots, y_n\} \tag{8.2}$$

$$Z = \{z_1, z_2, \ldots, z_n\} \tag{8.3}$$

An event is considered to be a recognizable symbol or property, e.g. the color red, within a particular modality.

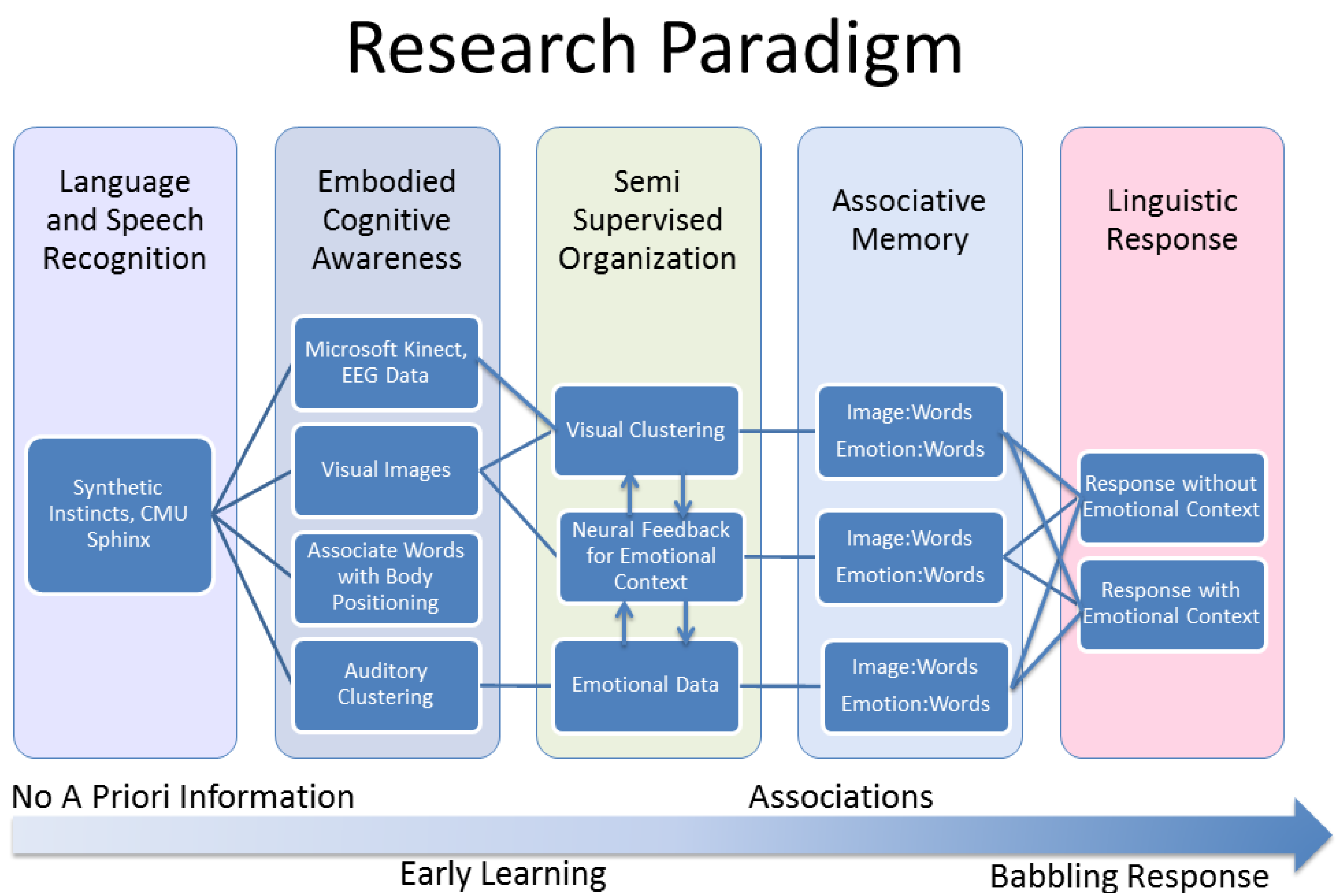


Figure 8.1: Experimental overview.

The purpose of the synthetic instinctual apparatus is to provide criteria for gathering this stream of information which can then be abstracted to simple vectors of "events". The synthetic instincts can be considered a filter and drive to some degree. This can be represented as a Kronecker comb function on the input streams.

$$\sum_{k=0}^{N} \delta\left[x - kN\right] \ where \ \delta_{ij} = \begin{cases} 0 & if \ i \neq j \\ 1 & if \ i = j \end{cases} \tag{8.4}$$

and we can then construct a matrix of this information.

The inner product of the input streams with the synthetic criteria are written as

$$x_i \cdot x_j^{'} = \sum_{ij} x_i \delta_{ij} x_j' \tag{8.5}$$

where x_i is the generic input stream and x' is the synthetic instinct criteria.

When data is received from these sensory vectors, the instinct increases the probability that any one of these data points will provide linguistically valuable concept data.

8.1.2 Cluster

Ideally we gather the information from these instincts into a triple (X, Y, Z) for concept construction purposes. We have connect our sensory inputs using a uniform temporal window with the assumption that is a reasonable basis for generating these triples.

A major component of this work also assumes that we do not need a priori information for these sensory streams. In the visual senses, self-organizing maps have proven to be effective at just such a task. In particular, SOMs also have a natural capacity for information compression which is also needed. Normally, temporal information is not overlayed within the SOM framework, this is addressed as we proceed in the concept building portion of this work.

We begin with a randomized space in $\mathbb{R}^2$ and where we provide the input from X, the visual stream directed by the visual synthetic instinct shown in Figure 8.2. This same formulation is used for all sensory streams in this research. We apply the standard Kohonen map algorithm to the information stream and follow the specification below:

$$W(t+1) = W(t) + \theta(t)\alpha(t)(X(t) - W(t)) \tag{8.6}$$

where $X(t)$ = input vector

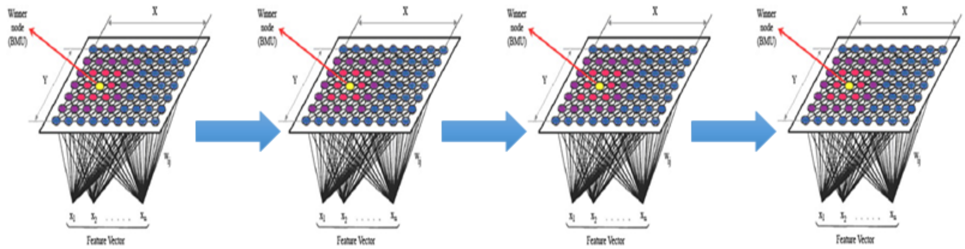


Figure 8.2: Storyboarding.

$W(t)$ = weight vector

$\alpha(t)$ = decreasing learning coefficient, $\alpha_0 e^{-\frac{t}{\lambda}}$

$\theta(t)$ = neighborhood function $\sigma_0 e^{-\frac{\left(\sqrt{\sum\limits_{i=0}^{n}(x_i-W_i)^2}\right)^2}{2\sigma^2(t)}}$ where σ_0 is the initial radius of the neighborhood

n = number of dimensions in the sample vector

The similarity between the data stream and nodes will be tested by using the minimum Euclidean distance, $\min_i \left(\sqrt{\sum\limits_{i=0}^{n}(x_i - W_i)^2}\right)$.

The next phase is called "Storyboarding", $\Sigma = \{s_1, s_2, \ldots, s_n\}$ is a sequence of SOMs in emotional order. These maps are generated during the time window and they provide information about a visual and spacial scene the android and human are experiencing.

We also have the auditory input stream Y being gathered and classified by a three-state hidden Markov model. The mathematical details can be seen in Section 4.2.2. This stream is conditioned by auditory synthetic instincts to focus on speech-like sounds. The Z component of the triple in this case refers to the emotion data collected by the EEG.

8.1.3 Swarm

The swarm enters as the information gathered from the streams assessed within a linguistic space which was developed from verbal interactions. The swarm uses a metric nearest-neighbor, $D(y_i, z_i)$, to determine the word-emotion proximity which provides context. This information is then used to help classify the interactions with the android and to suggest words for responses.

As the data streams are recorded, a distribution is associated with each stream.

8.1.4 Associative Memories

Once presented with an object several times and verbally communicated word(s) and sentences, we communicated words at random from the previous training session to the android to verify word-object association. This association can be described by the following: $S : (\Sigma \times Y) \rightarrow R$ where R is the response in this case. The response can be non-verbal; in this case it is the image associated in the memory of the android.

8.1.5 Reservoir Computing

The final stage takes all of the data from the previous linguistic interaction and attempts to learn from this data via a reservoir computing recurrent neural network. The result is also babbled, but training leads to an increased relevance for the words related to the verbal stimulus given and its emotional content.

Chapter 9

FINAL RESULTS AND DISCUSSION

No book can ever be finished. While working on it we learn just enough to find it immature the moment we turn away from it. - Karl Popper

9.1 Experiments

This section describes the experiments and the results.

9.1.1 Experiment 1: Associative Memory

The same objects used during the training sessions were used again to keep experimental confounds to a minimum. The visual objects are now associated with spoken word sounds. The iCub was taught for several runs and the association was tested by android's ability to recall an image that included the spoken object. The amount of focus time needed to be above a certain threshold as the synthetic instincts are running in parallel.

9.1.2 Results

The android recorded visual information as well auditory information in a database. This recording is associated with its interaction in real time. The system has the ability to listen and recognize speech, as mentioned in Section 5.1. We were able to see the image and the word(s) associated with it by prompting the android at a later time. To test the associative memory, several words that were uttered to the android earlier were repeated. They are expected to be in the lexicon, if the android recognized the word sounds during training. Next we waited to see the image that was presented by the android, to verify if it indeed recognized the word.

Figures 9.1 to 9.6 show the results of this experiment in the following format: Word Utterance, Image Retrieved.

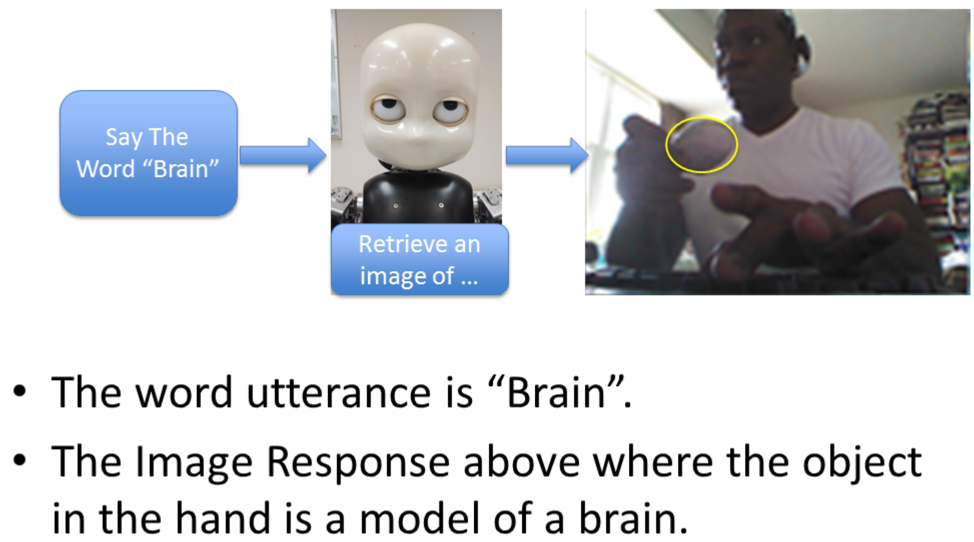


Figure 9.1: Associative memory response #1.

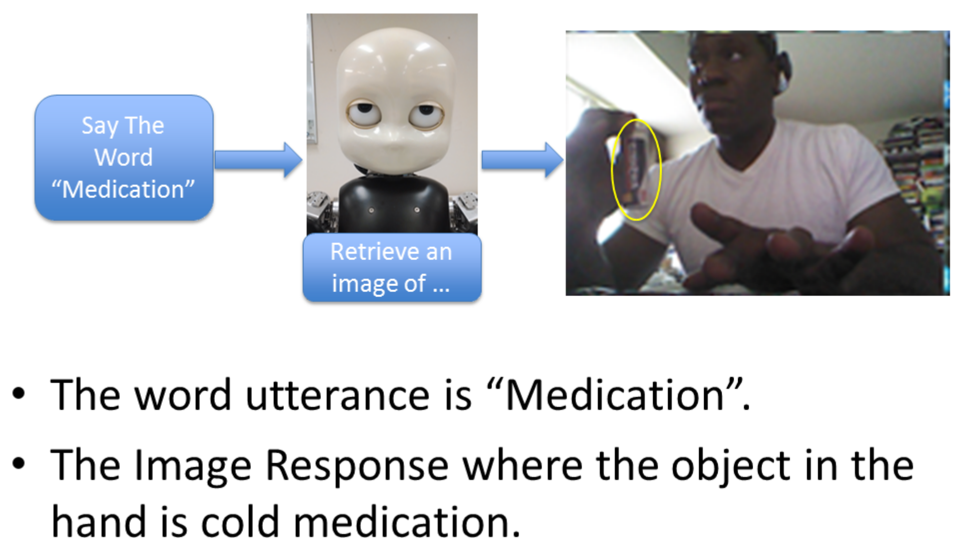


Figure 9.2: Associative memory response #2.

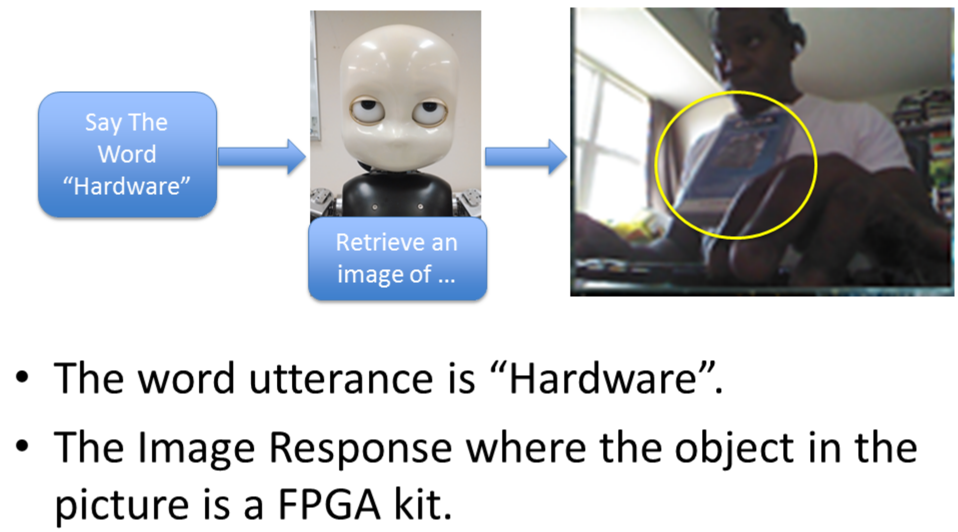


Figure 9.3: Associative memory response #3.

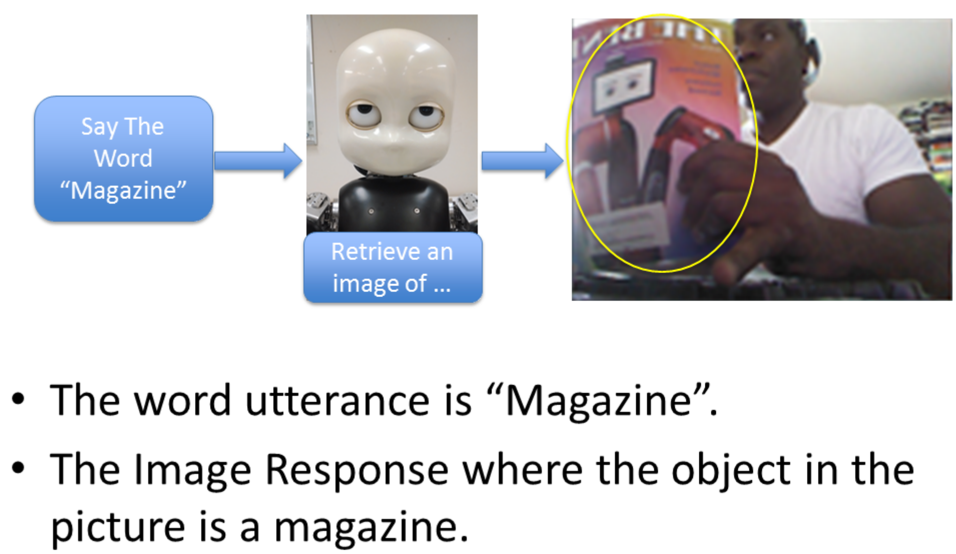


Figure 9.4: Associative memory response #4.

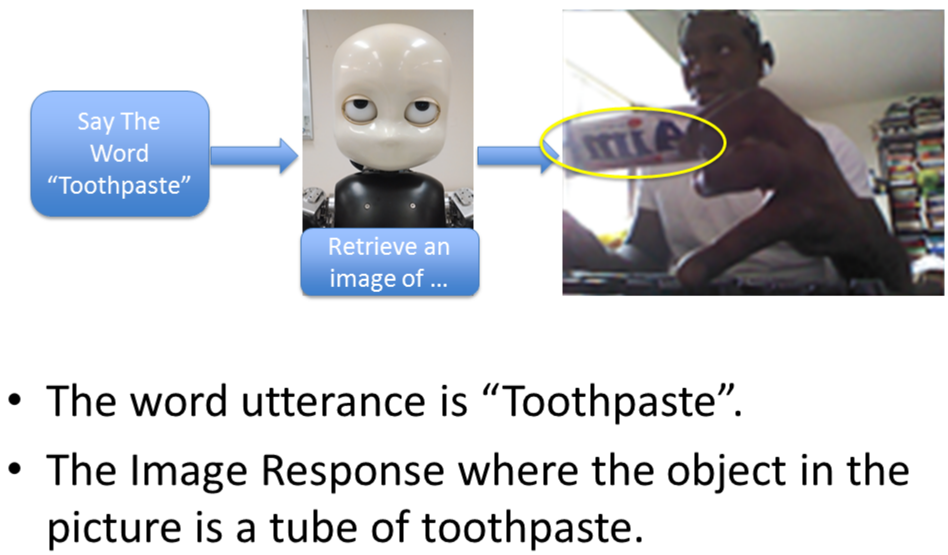


Figure 9.5: Associative memory response #5.

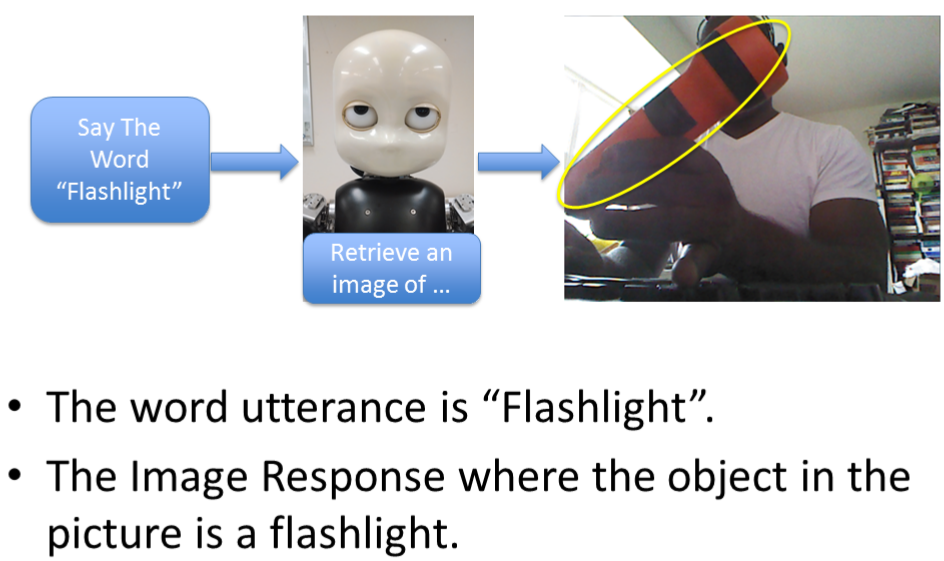


Figure 9.6: Associative memory response #6.

The mapping was not necessarily one-to-one which is the case in human memory as well.

The retrieval was successful, if the spoken utterance was recognized which happens about 90 percent of the trials. There are often several options available for the memory retrieval, for easy of utilization, the first image recorded on any trial was the image retrieved.

9.1.3 Experiment 2: Linguistic Babbling

In this experiment, several words are spoken to the robot for a training period. The android was given the chance to see the objects during the conversation and form the associative memory with the images. If the utterances were not recognized they were repeated until the correct word(s) were associated with the full visual field. In addition to auditory and visual information the android is simultaneously receiving real-time information about the following five emotional states via the EEG:

- Engagement: A feeling normally experienced as a level of alertness and/or the conscious direction of attention toward some task-relevant stimuli. It is often detected via increased physiological arousal and beta EEG wave forms in conjunction with a reduction in alpha wave expression.
- Excitement: An awareness or feeling of physiological arousal that can generally be described as positive. Excitement is characterized by activation in the sympathetic nervous system which results in a range of physiological responses including pupil dilation, eye widening, etc.
- Frustration: A reduction of engagement with an increase in excitement in this system. The physiological responses vary on the individual.
- Meditation: A suppression of excitement and frustration. A calming physiological response is also noted.
- Valence: A measure of the intrinsic attractiveness or aversivness caused by an event.

For proprietary reasons, Emotiv does not make the exact algorithms involved in these calculations public. The data is used to compare babbling driven with emotional context versus randomly selected babbling from the android.

9.1.4 Results

9.1.4.1 Speech Recognition Phase

This is the first phase of the experiment and Figures 9.7-9.9 show graphs of the words recognized by the iCub after training.

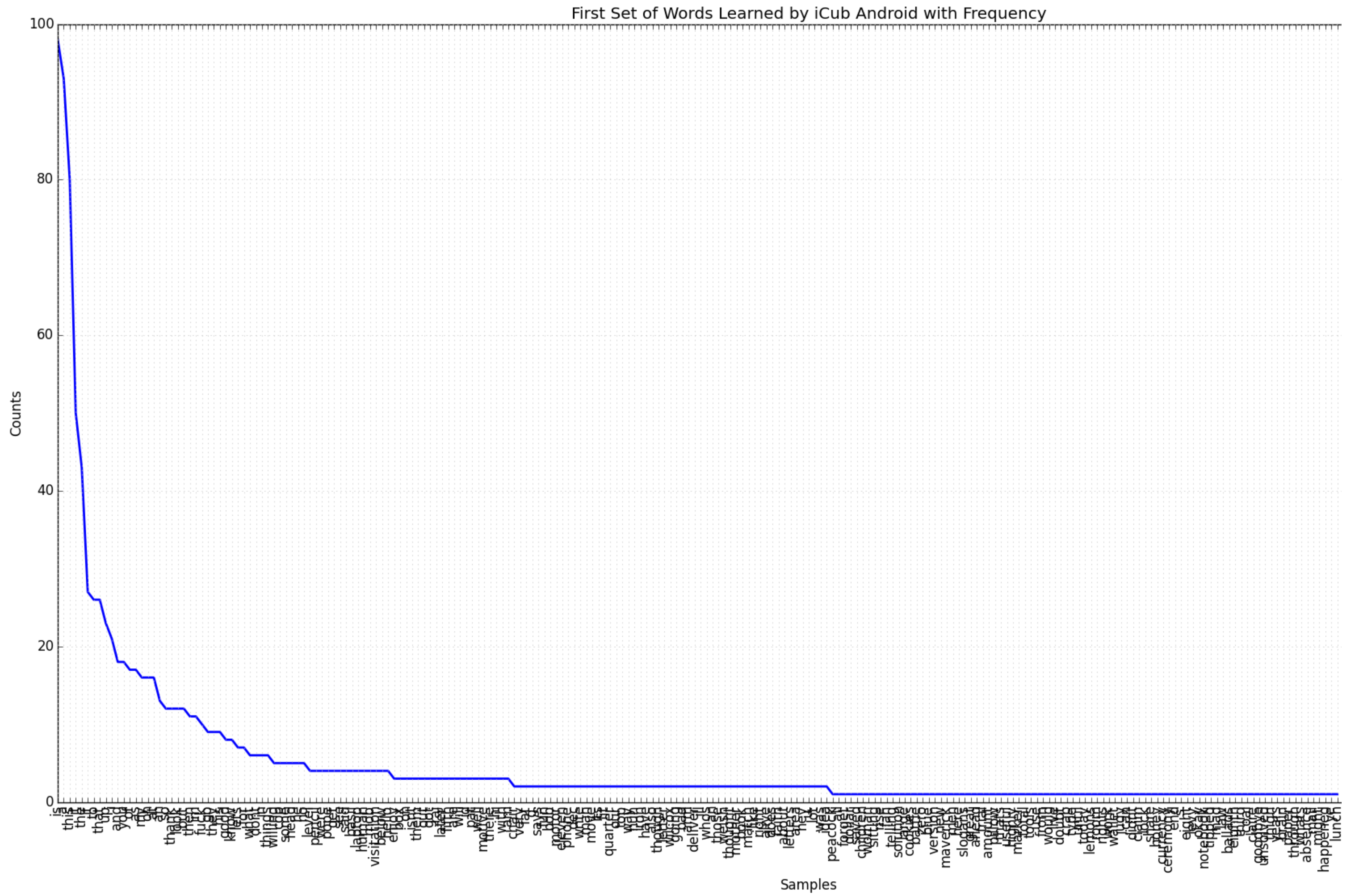


Figure 9.7: First set of iCub learned word counts.

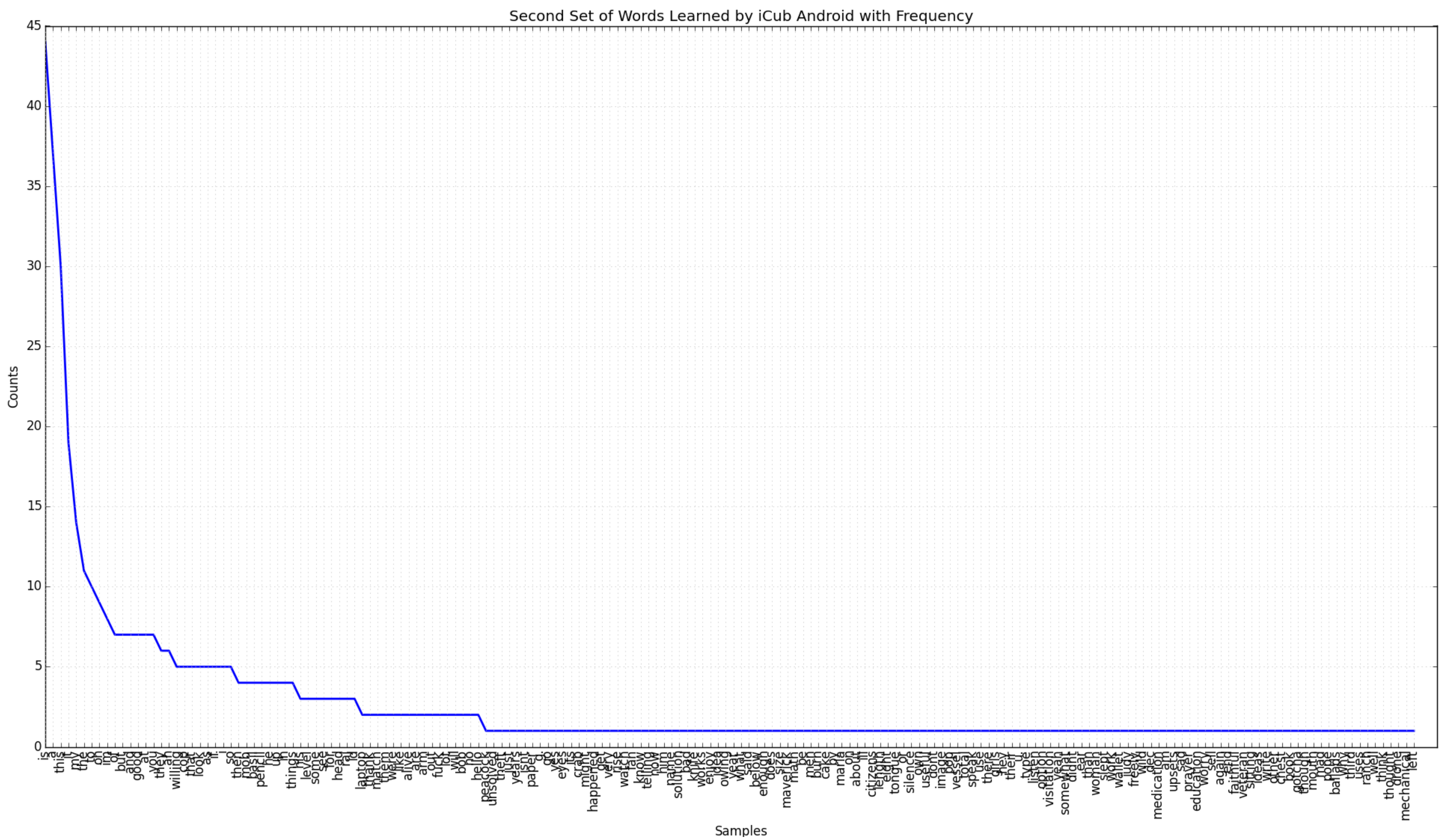


Figure 9.8: Second set of iCub learned words.

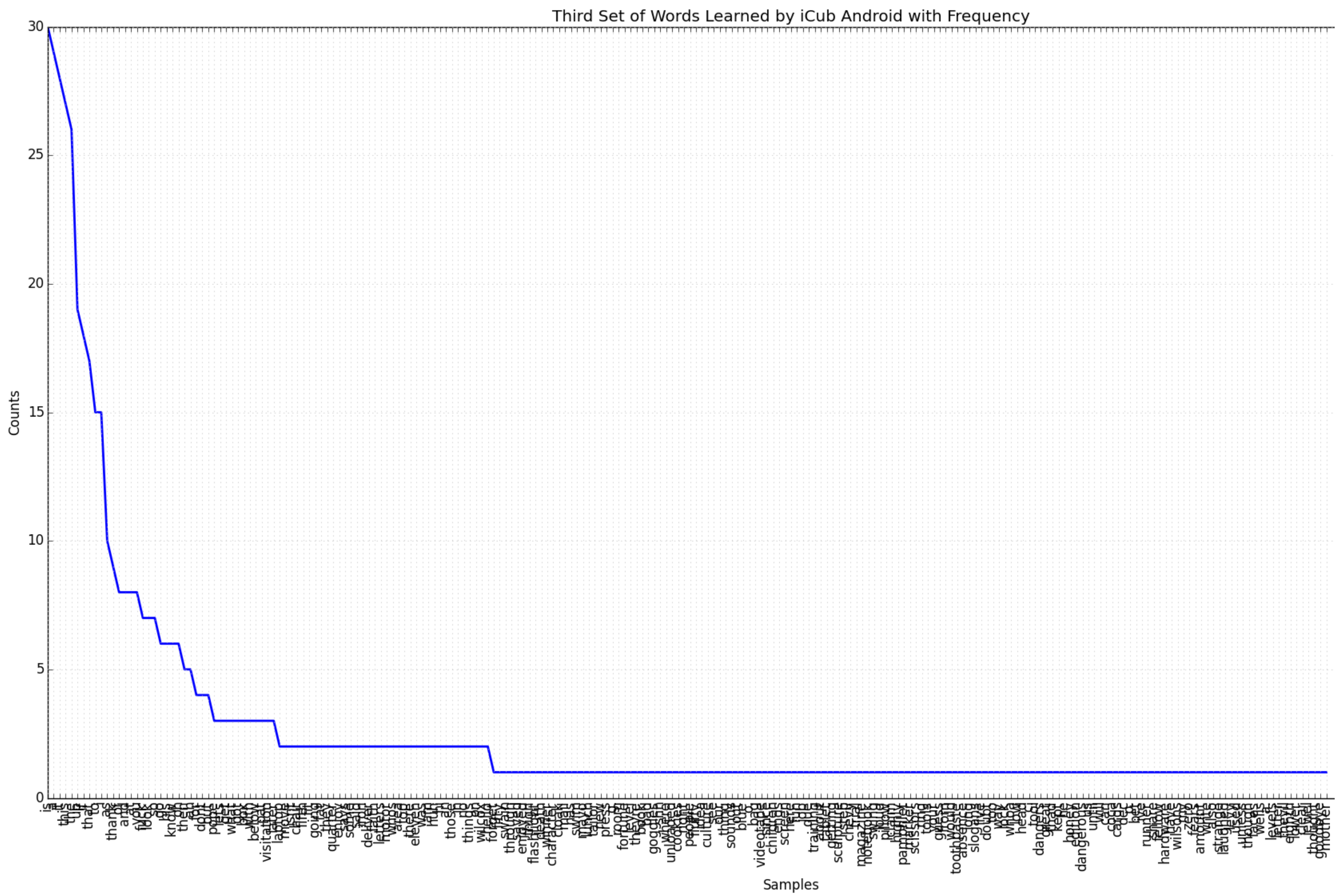


Figure 9.9: Third set of iCub learned words.

The iCub recognized the spoken utterance about 65-70% of the time of the utterances first presentation during the training trials. The percentage increased to approximately 90% (89.3%) with repeated utterance of the word and/or phrase. This unfortunately introduced misheard artifacts in to the iCub's lexicon, but this turned out to be useful since this phenomenon also occurs in human beings. The default auditory dictionary for phoneme recognition was used with pocketsphinx, a C version of the CMU Sphinx framework discussed in Section 5.1. By choosing the default dictionary, it makes the task of using this framework more consistent across users. The option to develop a more customized phoneme recognition lexicon was considered detrimental to the universality of this type of training.

9.1.4.2 Self-Organizing Mapping Phase

Figure 9.10 shows the results of the EEG data self-organized via a Kohonen map.

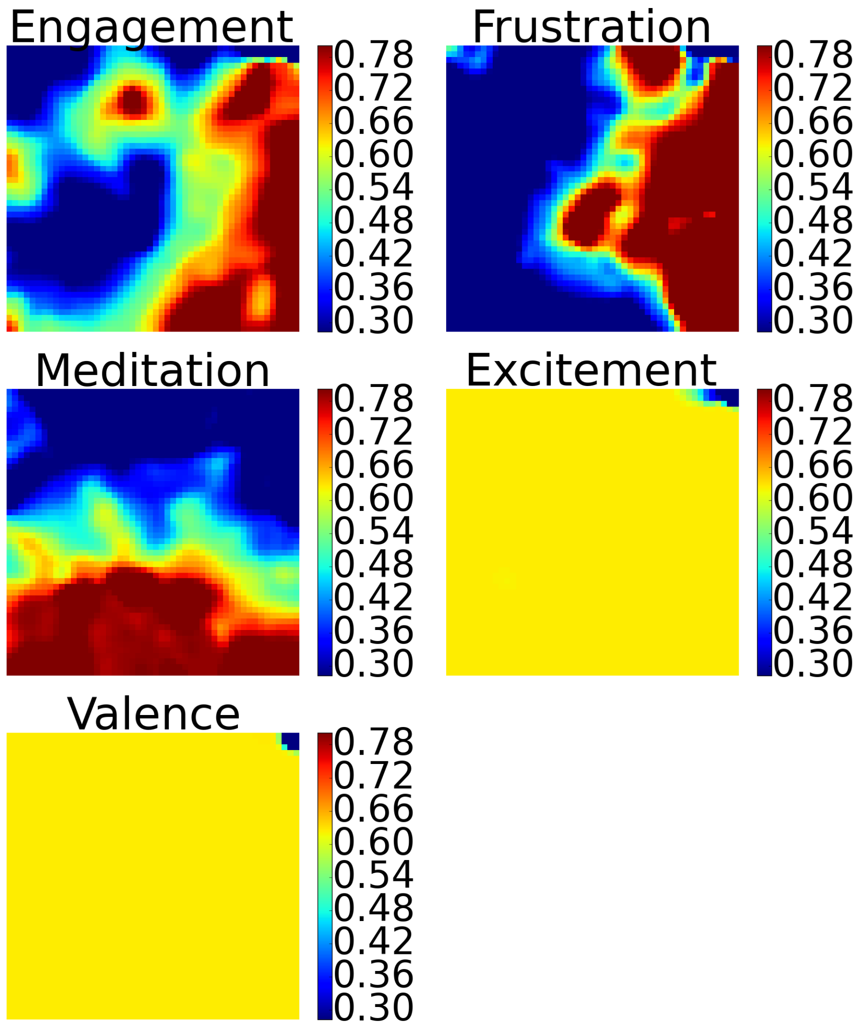


Figure 9.10: Kohonen maps for emotional states during live training.

The words and their corresponding emotional values from the EEG modality were checked for correlation across all five emotions. Each of the distinct strips corresponds to one of the emotional vectors, Engagement, Frustration, Meditation, Excitement and Valence. Figure 9.10 serves two purposes. First, it is a good check for emotional richness of the data-word collection process. If the subject was experiencing variations in his or her emotional response, one would expect that distribution of the correlation would follow a trend where the words are of a similar emotion and there should be roughly as many identifiable trends as measured emotional states. The second purpose is to verify that there is some overlap of the emotional content and the words as would be expected in normal language since all emotional levels are recorded for each word. We can observe both results in Figure 9.11.

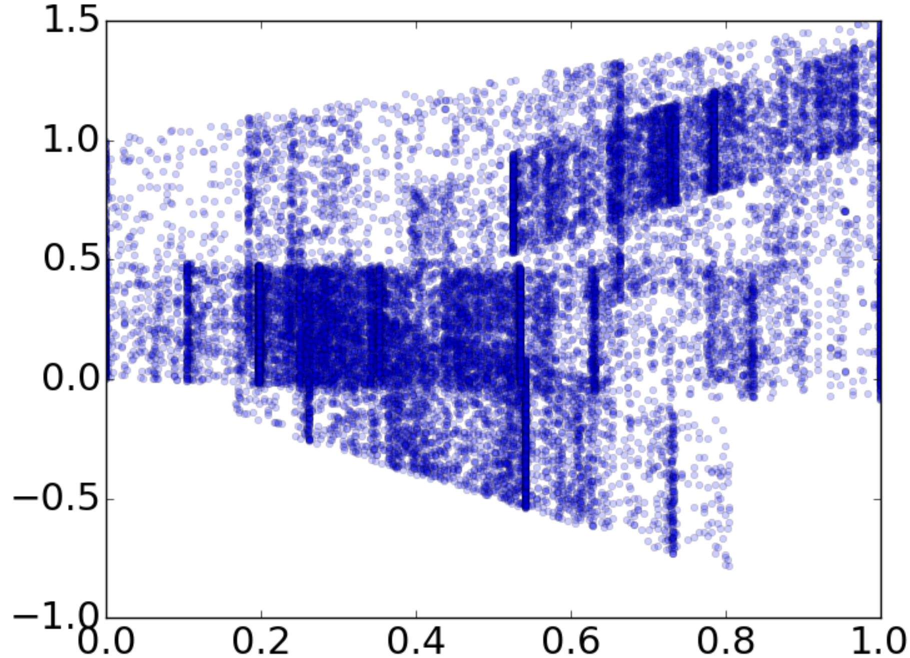


Figure 9.11: Word emotion correlation graph.

9.1.4.3 Swarm Phase

During this phase the emotional state information is provided to the swarm which then selects words from the iCub's lexicon that has been mapped to the SOM. The swarm follows a very simple set of rules:

1. Each member of the swarm can only have one word.
2. The word must have an emotional association within a specified range for the most dominant emotion at the measurement time.
3. The words are then passed to the reservoir computer in no particular order.

During the experiment the words selected were of the correct emotional range throughout the entire run (±5-10%). The parts of speech were not accounted for during the swarm

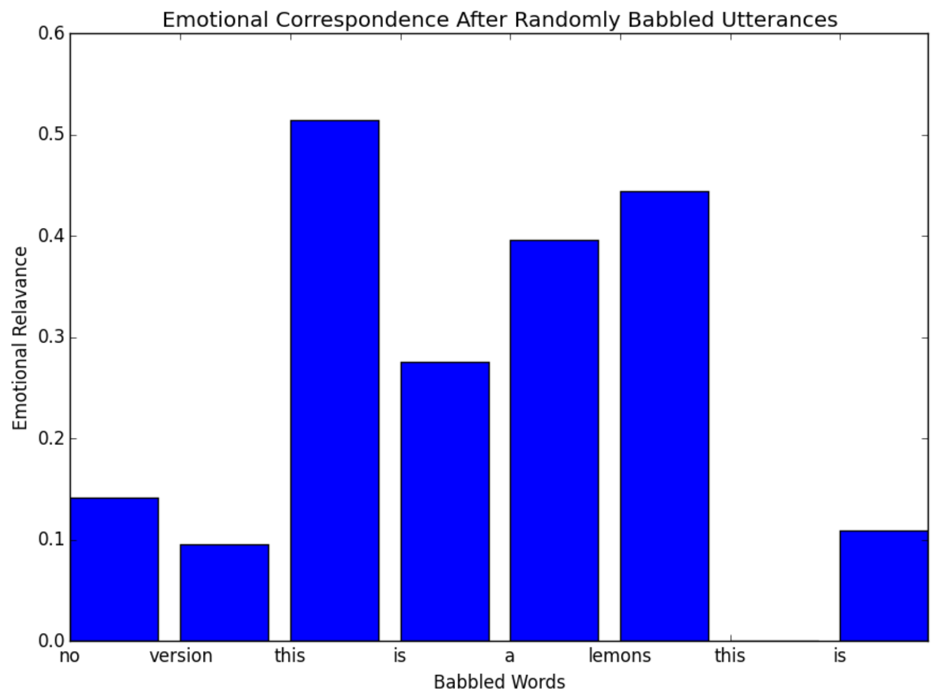


Figure 9.12: Sample of random babbling words.

collection process since babbling is a very early language interaction and it does not follow any real grammatical rules.

9.1.4.4 Reservoir Computing Phase

The reservoir computing phase allows for an attempt to make the babbling attempt more emotionally relevant. Figures 9.12 and 9.14 show the results from training the babbling on a reservoir computer of 100 neurons and access to 18 words at a time from the iCub's lexicon. The number of words had to be limited due to the computational complexity introduced with larger training sets and the limited computational abilities of the research machine. In cases where this number of neurons and/or words exceeded 100 or 18 respectively, computational overflow would occur in the machine which would cause the training to crash and immediately invalidating the session and responses.

9.1.4.5 Linguistic Babbling Phase

The android was allowed to babble randomly, seen in Figure 9.16, in response to emotion data and a spoken word. This was compared to the procedure described in the previous phase. The results are displayed in Figures 9.13 and 9.15.

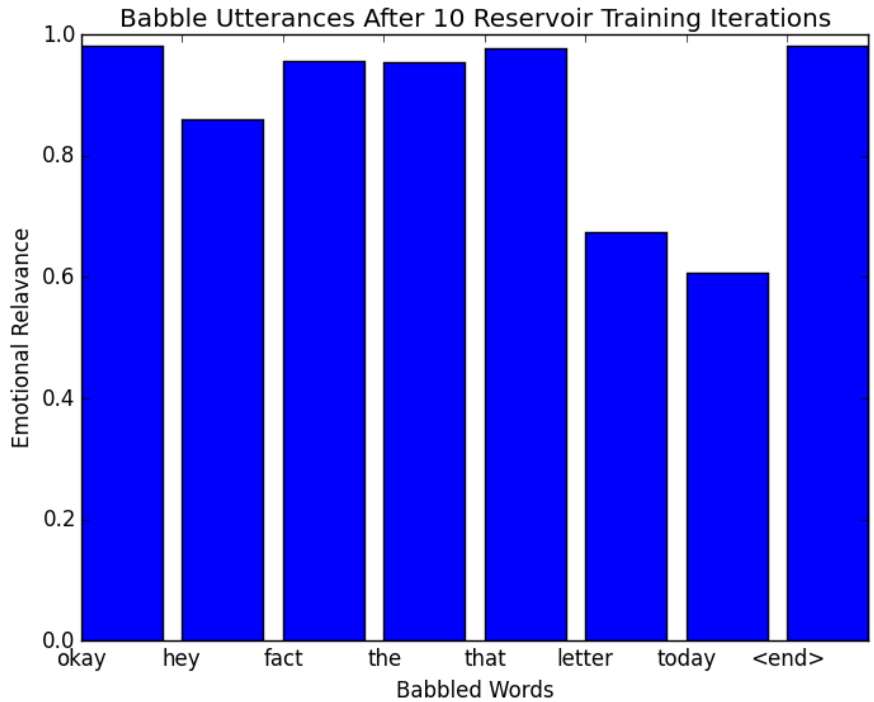


Figure 9.13: Sample of babbled words with emotional cues.

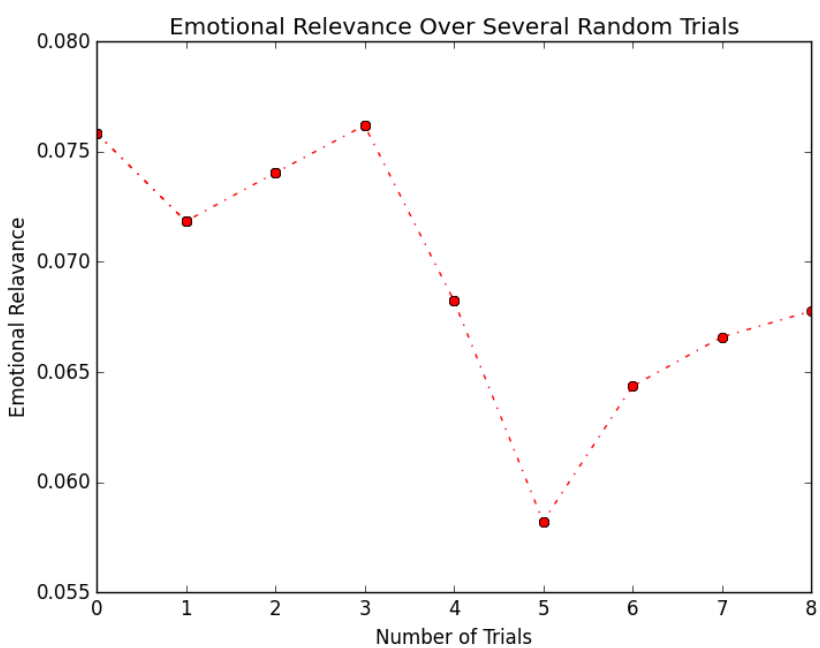


Figure 9.14: Randomly babbled words.

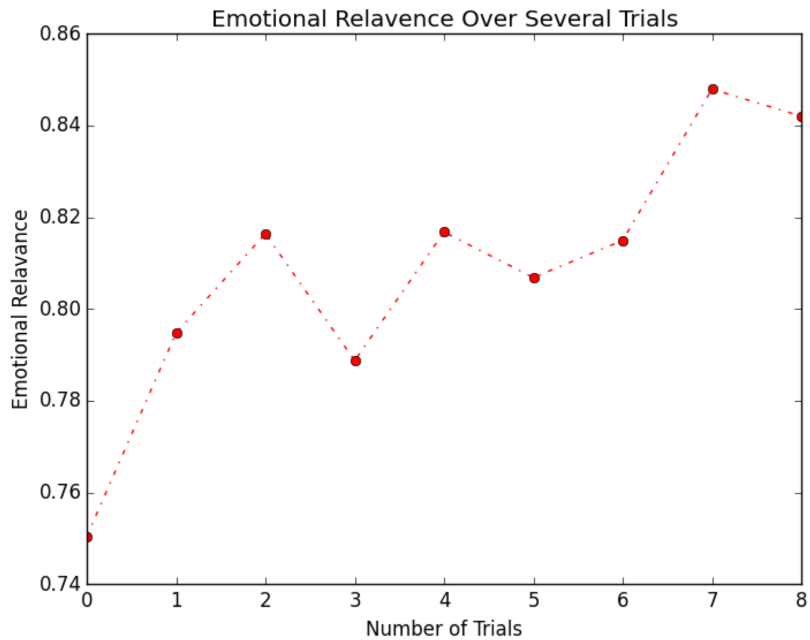


Figure 9.15: Babbled words from emotional cues.

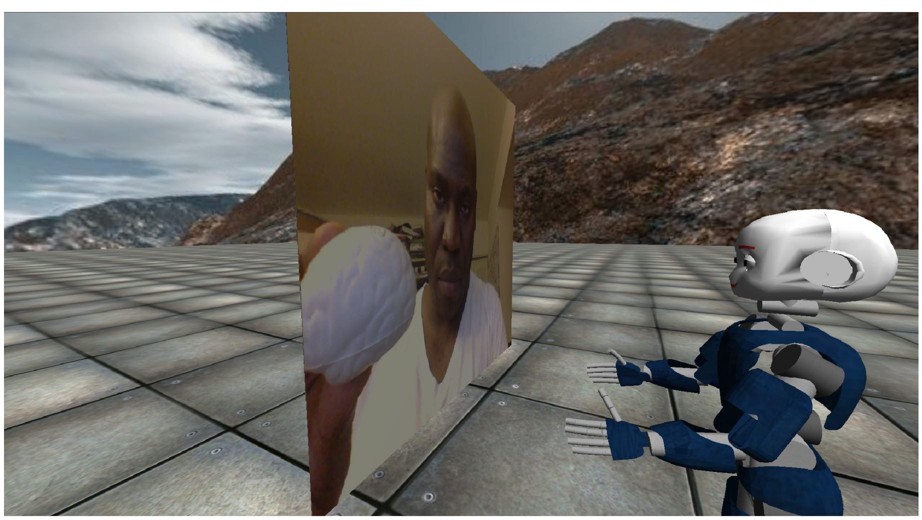


Figure 9.16: iCub verbally babbling from its virtual environment.

Chapter 10

DISCUSSION

All men are caught in an inescapable network of mutuality. – Dr. Martin Luther King Jr.

This work demonstrates a case where sensory fusion is a useful method for simple language acquisition. The results while not necessarily the likes of the systems on smart phones is actually many times more authentic in design, since not a single explicit language rule or instruction was used to generate demonstrably emotionally relevant responses with the babbling. We also see that Kohonen maps and reservoir computing are complementary in this regard as both methods can function without explicit supervision. The emergence of these emotionally tethered states also gives further credence to this approach. This phenomenon is interesting, particularly due to its similarities to behavior observed in newborn children [94, 95, 96, 97]. This may well imply that our linguistic response and acquisition process may have other hidden components, not previously considered. Further investigation could enable us to shine more light on this.

Another fact to keep in mind is that despite our inability to verify with 100% certainty the true correlates used by Emotiv for their emotional state classification, this research is not actually predicated on knowing the exact emotional states. The event of interest is the android synchronizing with any recorded emotional state from the experiment at the time of learning the words.

Chapter 11

FUTURE WORK

Had I the heavens embroidered cloths, Enwrought with golden and silver light, the blue and the dim and the dark cloths of night and light and the half light, I would spread the cloths under your feet: But I, being poor, have only my dreams; I have spread my dreams under your feet; Tread softly because you tread on my dreams. - W.B. Yeats, "He Wishes For the Cloths of Heaven"

11.1 Experimental Considerations

As this work proceeds, we should compare these results with a larger subject pool. This would allow for full sensory fusion and reduced sensory groups, e.g. audio or video removed from the fusion. Randomization of the emotion-word context would also help to remove unforseen biases. This has the added benefit of verifying this effect across subjects.

11.2 Motor-Linguistic Babbling and Response

Another dimension of this research would look at the ability to incorporate body language in the learning cycle as well. We have also provided the android with body positioning data, e.g. the real-time position of our head and arms while we were interacting with it verbally. This is an important layer of the early linguistic acquisition because approximately 55% of our language interaction is non-verbal, although this percentage can vary based on the specifics of the scenario [98, 99, 100, 101]. Since full language acquisition is the goal, it is of great importance to consider this modality further.

11.3 Neuromorphic Circuitry and the iCub

In this research, we were unable to fully utilize the neuromorphic work that was investigated earlier on, but this would provide another dimension to this type of research. Spiking neural networks are in many ways attempts to have direct biological models run on hardware. We could investigate what kind of emergence happens when the underlying neural network is based on a truly biological phenomenon.

11.4 The iCub's Computer Vision Capabilities and Associative Memories

If the iCub's computer vision capabilities were improved to allow for individual object and pattern recognition, especially via this kind of paradigm, one could imagine a similar experiment being conducted with visual states versus verbal language states.

REFERENCES

[1] K. Gödel, *On formally undecidable propositions of Principia mathematica and related systems.* New York: Dover Publications, 1992.

[2] B. Goertzel and C. Pennachin, Eds., *Artificial general intelligence.* Springer, 2007, vol. 2.

[3] M. Hutter, "A theory of universal artificial intelligence based on algorithmic complexity," *arXiv preprint cs/0004001*, 2000.

[4] S. Legg and M. Hutter, "Universal intelligence: A definition of machine intelligence," *Minds and Machines*, vol. 17, no. 4, pp. 391–444, 2007.

[5] D. Poole, A. Mackworth, and R. Goebel, "Computational intelligence and knowledge," *Computational intelligence: a logical approach*, pp. 1–22, 1998.

[6] L. S. Gottfredson, "Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography," *Intelligence*, vol. 24, no. 1, pp. 13–23, 1997.

[7] U. Neisser, G. Boodoo, T. J. Bouchard Jr, A. W. Boykin, N. Brody, S. J. Ceci, D. F. Halpern, J. C. Loehlin, R. Perloff, R. J. Sternberg et al., "Intelligence: knowns and unknowns,," *American psychologist*, vol. 51, no. 2, p. 77, 1996.

[8] S. Levinson, *Mathematical models for speech technology.* John Wiley & Sons, 2005.

[9] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience.* Lippincott Williams & Wilkins, 2007, vol. 2.

[10] J. Arnold, *Affect in language learning.* Ernst Klett Sprachen, 1999.

[11] P. M. Niedenthal, "Embodying emotion," *Science*, vol. 316, no. 5827, pp. 1002–1005, 2007.

[12] P. MacIntyre and T. Gregersen, "Emotions that facilitate language learning: The positive-broadening power of the imagination," *Studies in Second Language Learning and Teaching*, no. II-2, pp. 193–213, 2012.

[13] N. Wiener, *Cybernetics: Or the Control and Communication in the Animal and the Machine.* Martino Fine Books, Oct. 2013.

[14] A. M. Turing, “On computable numbers, with an application to the entscheidungsproblem,” *J. of Math*, vol. 58, pp. 345–363, 1936.

[15] A. M. Turing, “Computing machinery and intelligence,” *Mind*, pp. 433–460, 1950.

[16] K. Lorenz, *Behind The Mirror: A Search for a Natural History of Human Knowledge*. New York: Mariner Books, Sep. 1978.

[17] K. M. Haroon, “The neurochip: A new multielectrode device for stimulating and recording from cultured neurons,” *Biophysical Journal*, vol. 106, no. 2, p. 793a, 2014.

[18] B. E. Swartz, “The advantages of digital over analog recording techniques,” *Electroencephalography and clinical neurophysiology*, vol. 106, no. 2, pp. 113–117, 1998.

[19] B. Grözinger, H. Kornhuber, and J. Kriebel, “Methodological problems in the investigation of cerebral potentials preceding speech: determining the onset and suppressing artefacts caused by speech,” *Neuropsychologia*, vol. 13, no. 3, pp. 263–270, 1975.

[20] L. Y. Ganushchak, I. K. Christoffels, and N. O. Schiller, “The use of electroencephalography in language production research: a review,” *Frontiers in psychology*, vol. 2, 2011.

[21] C. Moonen, P. van Zijl, J. A. Frank, D. Le Bihan, and E. D. Becker, “Functional magnetic resonance imaging in medicine and physiology,” *Science*, vol. 250, no. 4977, pp. 53–61, 1990. [Online]. Available: http://www.sciencemag.org/content/250/4977/53.short

[22] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional magnetic resonance imaging*. Sinauer Associates Sunderland, MA, 2004.

[23] A. C. Grant, S. G. Abdel-Baki, J. Weedon, V. Arnedo, G. Chari, E. Koziorynska, C. Lushbough, D. Maus, T. McSween, K. A. Mortati et al., “EEG interpretation reliability and interpreter confidence: A large single-center study,” *Epilepsy & Behavior*, vol. 32, pp. 102–107, 2014.

[24] O. Sporns, “The human connectome: a complex network,” *Annals of the New York Academy of Sciences*, vol. 1224, no. 1, pp. 109–125, 2011. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2010.05888.x/full

[25] L. Cammoun, X. Gigandet, D. Meskaldji, J. P. Thiran, O. Sporns, K. Q. Do, P. Maeder, R. Meuli, and P. Hagmann, “Mapping the human connectome at multiple scales with diffusion spectrum mri,” *Journal of neuroscience methods*, vol. 203, no. 2, pp. 386–397, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165027011005991

[26] M. Axer, K. Amunts, D. Grässel, C. Palm, J. Dammers, H. Axer, U. Pietrzyk, and K. Zilles, “A novel approach to the human connectome: ultra-high resolution mapping of fiber tracts in the brain,” *Neuroimage*, vol. 54, no. 2, pp. 1091–1101, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S105381191001178X

[27] A. L. Hodgkin and A. F. Huxley, “A quantitative description of membrane current and its application to conduction and excitation in nerve,” *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, Aug. 1952, 15132 PMID: 12991237. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1392413/

[28] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” *The bulletin of mathematical biophysics*, vol. 5, no. 4, pp. 115–133, 1943. [Online]. Available: http://link.springer.com/article/10.1007/BF02478259

[29] P. Werbos, “Beyond regression: New tools for prediction and analysis in the behavioral sciences,” Ph.D. dissertation, Harvard University, 1974.

[30] T. Kohonen, “The self-organizing map,” *Neurocomputing*, vol. 21, no. 13, pp. 1–6, Nov. 1998, 00665. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231298000307

[31] T. Rumbell, S. L. Denham, and T. Wennekers, “A spiking self-organizing map combining stdp, oscillations, and continuous learning,” *IEEE Transactions on Neural Networks*, 2014.

[32] Z. N. B. Salem, F. Mouria-beji, and F. Kamoun, “Spatio-temporal organization map: A speech recognition application,” in *Artificial Neural Networks: Biological Inspirations–ICANN 2005.* Springer, 2005, pp. 371–378.

[33] N. R. Euliano and J. C. Principe, “Spatio-temporal self-organizing feature maps,” in *Neural Networks, 1996., IEEE International Conference on*, vol. 4. IEEE, 1996, pp. 1900–1905.

[34] K. Horio and T. Yamakawa, “Feedback self-organizing map and its application to spatio-temporal pattern classification,” *International Journal of Computational Intelligence and Applications*, vol. 1, no. 01, pp. 1–18, 2001.

[35] N. R. Euliano and J. C. Principe, “A spatio-temporal memory based on soms with activity diffusion,” *Kohonen maps*, pp. 253–266, 1999.

[36] P. F. Dominey, “Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning,” *Biological cybernetics*, vol. 73, no. 3, pp. 265–274, 1995. [Online]. Available: http://link.springer.com/article/10.1007/BF00201428

[37] T. Natschläger, W. Maass, and H. Markram, “The" liquid computer": A novel strategy for real-time computing on time series,” *Special issue on Foundations of Information Processing of TELEMATIK*, vol. 8, pp. 39–43, 2002. [Online]. Available: http://infoscience.epfl.ch/record/117806

[38] H. Jaeger, “The "echo state" approach to analysing and training recurrent neural networks-with an erratum note,” *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, p. 34, 2001.

[39] J. Steil, “Backpropagation-decorrelation: online recurrent learning with o(n) complexity,” in *2004 IEEE International Joint Conference on Neural Networks, 2004. Proceedings*, vol. 2, July 2004, 00139. pp. 843–848 vol.2.

[40] J. Szary, B. Kerster, and C. Kello, “What makes a brain smart? reservoir computing as an approach for general intelligence,” in *Artificial General Intelligence*, ser. Lecture Notes in Computer Science, J. Schmidhuber, K. R. Thórisson, and M. Looks, Eds. Springer Berlin Heidelberg, Jan. 2011, no. 6830, pp. 407–413, 00000.

[41] P. F. Dominey, M. Hoen, J.-M. Blanc, and T. Lelekov-Boissard, “Neurological basis of language and sequential cognition: evidence from simulation, aphasia, and erp studies,” *Brain and language*, vol. 86, no. 2, pp. 207–225, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0093934X02005291

[42] P. F. Dominey, M. Hoen, and T. Inui, “A neurolinguistic model of grammatical construction processing,” *Journal of Cognitive Neuroscience*, vol. 18, no. 12, pp. 2088–2107, 2006. [Online]. Available: http://www.mitpressjournals.org/doi/abs/10.1162/jocn.2006.18.12.2088

[43] P. F. Dominey and F. Ramus, “Neural network processing of natural language: I. sensitivity to serial, temporal and abstract structure of language in the infant,” *Language and Cognitive Processes*, vol. 15, no. 1, pp. 87–127, 2000. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/016909600386129

[44] H. Jaeger, M. Lukoševičius, D. Popovici, and U. Siewert, “Optimization and applications of echo state networks with leaky-integrator neurons,” *Neural Networks*, vol. 20, no. 3, pp. 335–352, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S089360800700041X

[45] H. Jaeger and H. Haas, “Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication,” *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004, 00816 PMID: 15064413. [Online]. Available: http://www.sciencemag.org/content/304/5667/78

[46] H. Jaeger, “Tutorial on training recurrent neural networks, covering BPTT, RTRL, EKF and the "echo state network" approach, german nat,” *Res. Inst. Inform. Technol., Sankt Augustin, Germany, Rep*, vol. 159, 2002.

[47] W. Maass, T. Natschläger, and H. Markram, “Real-time computing without stable states: A new framework for neural computation based on perturbations,” *Neural computation*, vol. 14, no. 11, pp. 2531–2560, 2002.

[48] W. Maass, T. Natschlager, and H. Markram, “A model for real-time computation in generic neural microcircuits,” *Advances in neural information processing systems*, pp. 229–236, 2003.

[49] T. Natschläger, H. Markram, and W. Maass, "Computer models and analysis tools for neural microcircuits," in *Neuroscience Databases.* Springer, 2003, pp. 123–138.

[50] W. Maass, P. Joshi, and E. D. Sontag, "Principles of real-time computing with feedback applied to cortical microcircuit models," *Advances in neural information processing systems*, vol. 18, p. 835, 2006.

[51] A. F. Atiya and A. G. Parlos, "New results on recurrent network training: unifying the algorithms and accelerating convergence," *Neural Networks, IEEE Transactions on*, vol. 11, no. 3, pp. 697–709, 2000.

[52] M. Lafourcade, F. Guinand et al., "Ants for natural language processing," 2005.

[53] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, Apr. 1982. [Online]. Available: http://www.pnas.org/content/79/8/2554

[54] E. Fredkin, "An informational process based on reversible universal cellular automata," *Physica D: Nonlinear Phenomena*, vol. 45, no. 1, pp. 254–270, 1990. [Online]. Available: http://www.sciencedirect.com/science/article/pii/016727899090186S

[55] E. Domany and W. Kinzel, "Equivalence of cellular automata to ising models and directed percolation," *Physical review letters*, vol. 53, no. 4, p. 311, 1984. [Online]. Available: http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.53.311

[56] G. Grinstein, C. Jayaprakash, and Y. He, "Statistical mechanics of probabilistic cellular automata," *Physical review letters*, vol. 55, no. 23, p. 2527, 1985. [Online]. Available: http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.55.2527

[57] T. E. Ingerson and R. L. Buvel, "Structure in asynchronous cellular automata," *Physica D: Nonlinear Phenomena*, vol. 10, no. 1, pp. 59–68, 1984. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0167278984902495

[58] J. L. Lebowitz, C. Maes, and E. R. Speer, "Statistical mechanics of probabilistic cellular automata," *Journal of Statistical Physics*, vol. 59, no. 1-2, pp. 117–170, 1990. [Online]. Available: http://link.springer.com/article/10.1007/BF01015566

[59] S. Wolfram, "Statistical mechanics of cellular automata," *Reviews of modern physics*, vol. 55, no. 3, p. 601, 1983. [Online]. Available: http://journals.aps.org/rmp/abstract/10.1103/RevModPhys.55.601

[60] S. Wolfram, "Universality and complexity in cellular automata," *Physica D: Nonlinear Phenomena*, vol. 10, no. 1, pp. 1–35, 1984. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0167278984902458

[61] J. Haugeland, Ed., *Semantic engines: An introduction to mind design.* MIT Press, 1981.

[62] K. Squire and S. Levinson, “HMM-based concept learning for a mobile robot,” *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 2, pp. 199 –212, Apr. 2007.

[63] B. Szatmáry and E. M. Izhikevich, “Spike-timing theory of working memory,” *PLoS Comput Biol*, vol. 6, no. 8, p. e1000879, Aug. 2010. [Online]. Available: http://dx.doi.org/10.1371/journal.pcbi.1000879

[64] N. Watier and C. Collin, “The effects of distinctiveness on memory and metamemory for face–name associations,” *Memory*, vol. 20, no. 1, pp. 73–88, 2012. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/09658211.2011.637935

[65] M. Gazzaniga, R. Ivry, and G. Mangun, “Learning and memory,” *Cognitive neuroscience: The biology of the mind*, pp. 312–363, 2009.

[66] W. Timberlake, “Behavior systems, associationism, and pavlovian conditioning,” *Psychonomic Bulletin & Review*, vol. 1, no. 4, pp. 405–420, 1994. [Online]. Available: http://link.springer.com/article/10.3758/BF03210945

[67] W. Timberlake, “An animal-centered, causal-system approach to the understanding and control of behavior,” *Applied Animal Behaviour Science*, vol. 53, no. 1, pp. 107–129, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168159196011549

[68] S. J. Shettleworth, *Cognition, evolution, and behavior*. Oxford University Press, 2009.

[69] D. Saumier and H. Chertkow, “Semantic memory,” *Current neurology and neuroscience reports*, vol. 2, no. 6, pp. 516–522, 2002. [Online]. Available: http://link.springer.com/article/10.1007/s11910-002-0039-9

[70] C. Mead, *Analog VLSI and Neural Systems*, 1st ed. Addison Wesley Publishing Company, Jan. 1989.

[71] C. Mead, “Neuromorphic electronic systems,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1629 –1636, Oct. 1990, 00425.

[72] K. Boahen, “Neuromorphic microchips,” *Scientific American*, vol. 292, no. 5, pp. 56–63, May 2005.

[73] R. Silver, K. Boahen, S. Grillner, N. Kopell, and K. L. Olsen, “Neurotech for neuroscience: unifying concepts, organizing principles, and emerging tools,” *The Journal of Neuroscience*, vol. 27, no. 44, pp. 11 807–11 819, 2007. [Online]. Available: http://www.jneurosci.org/content/27/44/11807.short

[74] “Neurogrid,” June 2015, page Version ID: 668763877. [Online]. Available: https://en.wikipedia.org/w/index.php?title=NeurogridI&oldid=668763877

[75] R. Ananthanarayanan, S. K. Esser, H. D. Simon, and D. S. Modha, "The cat is out of the bag: cortical simulations with 10^9 neurons, 10^{13} synapses," in *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*, ser. SC '09. New York, NY, USA: ACM, 2009. [Online]. Available: http://doi.acm.org/10.1145/1654059.1654124 pp. 63:1–63:12.

[76] D. S. Modha, R. Ananthanarayanan, S. K. Esser, A. Ndirango, A. J. Sherbondy, and R. Singh, "Cognitive computing," *Communications of the ACM*, vol. 54, no. 8, p. 62, Aug. 2011. [Online]. Available: http://cacm.acm.org/magazines/2011/8/114944-cognitive-computing

[77] J. Seo, B. Brezzo, Y. Liu, B. D. Parker, S. K. Esser, R. K. Montoye, B. Rajendran, J. A. Tierno, L. Chang, D. S. Modha, and D. J. Friedman, "A 45nm CMOS neuromorphic chip with a scalable architecture for learning in networks of spiking neurons," in *2011 IEEE Custom Integrated Circuits Conference (CICC)*. IEEE, Sep. 2011, pp. 1–4.

[78] A. Cassidy, S. Denham, P. Kanold, and A. Andreou, "FPGA based silicon spiking neural array," in *Biomedical Circuits and Systems Conference, 2007. BIOCAS 2007. IEEE*, Nov. 2007, 00025. pp. 75 –78.

[79] H. Shayani, P. Bentley, and A. Tyrrell, "Hardware implementation of a bio-plausible neuron model for evolution and growth of spiking neural networks on FPGA," in *NASA/ESA Conference on Adaptive Hardware and Systems, 2008. AHS '08*, June 2008, pp. 236 –243.

[80] M. Pearson, A. Pipe, B. Mitchinson, K. Gurney, C. Melhuish, I. Gilhespy, and M. Nibouche, "Implementing spiking neural networks for real-time signal-processing and control applications: A model-validated FPGA approach," *Neural Networks, IEEE Transactions on*, vol. 18, no. 5, pp. 1472 –1487, Sep. 2007.

[81] J. Li, Y. Katori, and T. Kohno, "An FPGA-based silicon neuronal network with selectable excitability silicon neurons," *Frontiers in Neuromorphic Engineering*, vol. 6, p. 183, 2012.

[82] A. S. Cassidy, J. Georgiou, and A. G. Andreou, "Design of silicon brains in the nano-CMOS era: Spiking neurons, learning synapses and neural architecture optimization," *Neural Networks*, vol. 45, pp. 4–26, Sep. 2013, 00004. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608013001597

[83] A. Cassidy, A. G. Andreou, and J. Georgiou, "Design of a one million neuron single FPGA neuromorphic system for real-time multimodal scene analysis," in *2011 45th Annual Conference on Information Sciences and Systems (CISS)*. IEEE, Mar. 2011, pp. 1–6.

[84] G. Indiveri, R. Douglas, and L. Smith, "Silicon neurons," *Scholarpedia*, vol. 3, no. 3, p. 1887, 2008.

[85] G. Deboeck and T. Kohonen, *Visual explorations in finance: with self-organizing maps.* Springer Science & Business Media, 2013.

[86] C.-S. Poon and K. Zhou, “Neuromorphic silicon neurons and large-scale neural networks: Challenges and opportunities,” *Frontiers in Neuroscience*, vol. 5, Sep. 2011.

[87] P. Sharma, “FPGA implementation of artificial neural networks,” BTech, 2007. [Online]. Available: http://ethesis.nitrkl.ac.in/4217/

[88] G. Indiveri, B. Linares-Barranco, T. J. Hamilton, R. Etienne-Cummings, T. Delbruck, S.-C. Liu, P. Häfliger, S. Renaud, J. Schemmel, G. Cauwenberghs, J. Arthur, S. Saighi, J. Wijekoon, and K. Boahen, “Neuromorphic silicon neuron circuits,” *Frontiers in Neuromorphic Engineering*, vol. 5, p. 73, 2011, 00129.

[89] V. Salapura, M. Gschwind, and O. Maischberger, “A fast FPGA implementation of a general purpose neuron,” in *Proceedings of the 4th International Workshop on Field-Programmable Logic and Applications: Field-Programmable Logic, Architectures, Synthesis and Applications*, ser. FPL ’94. London, UK, UK: Springer-Verlag, 1994. [Online]. Available: http://dl.acm.org/citation.cfm?id=647921.740703 pp. 175–182.

[90] A. R. Omondi and J. C. Rajapakse, *FPGA implementations of neural networks.* Springer, 2006, vol. 365.

[91] L. Rabiner, “A tutorial on hidden markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[92] P. Placeway, S. Chen, M. Eskenazi, U. Jain, V. Parikh, B. Raj, M. Ravishankar, R. Rosenfeld, K. Seymore, M. Siegler et al., “The 1996 hub-4 sphinx-3 system,” in *Proc. DARPA Speech recognition workshop.* Citeseer, 1997, pp. 85–89.

[93] R. A. Zwaan, “Embodiment and language comprehension: reframing the discussion,” *Trends in cognitive sciences*, vol. 18, no. 5, pp. 229–234, 2014.

[94] H. Kellerman, “An epigenetic theory of emotions in early development,” *Emotion: Theory research and experience*, vol. 2, pp. 315–349, 1983.

[95] R. Plutchik, “Emotions in early development: A psychoevolutionary approach,” *Emotion: Theory, research, and experience*, vol. 2, pp. 221–257, 1983.

[96] R. Plutchik and H. Kellerman, *Emotions in early development.* Academic Press, 2013, vol. 2.

[97] C. Trevarthen, “The function of emotions in early infant communication and development,” *New perspectives in early communicative development*, vol. 3, pp. 48–81, 1993.

[98] J. K. Burgoon, L. K. Guerrero, and K. Floyd, *Nonverbal communication.* Allyn & Bacon Boston, MA, 2010.

[99] R. G. Harper, A. N. Wiens, and J. D. Matarazzo, *Nonverbal communication: The state of the art.* John Wiley & Sons, 1978.

[100] M. Knapp, J. Hall, and T. Horgan, *Nonverbal communication in human interaction.* Cengage Learning, 2013.

[101] A. Mehrabian, *Nonverbal communication.* Transaction Publishers, 1977.