The Brain Sciences Journal
The Official Journal of The Brain Sciences Foundation

Editorial Board

Editors-in-Chief
Dr. Richard Wirt – Intel Corporation, Brain Sciences Senior Fellow

Associate Editors
Prof. Shlomo Argamon – Illinois Institute of Technology
Dr. Jeroen Bergmann – Imperial College London
Dr. Erik Cambria – National University of Singapore
Prof. Martine Collard – University of the French West Indies and Guiana
Prof. Tariq Durrani – University of Strathclyde
Mr. Sergey Kanareykin – Make A Mind Co.
Dr. Michael Kohl – Cambridge University
Prof. Mark Moss – Boston University
Prof. Tomaso Poggio – Massachusetts Institute of Technology
Prof. Colette Roland – La Sorbonne University

Editorial Staff
Mr. Mark Atallah – MIT Mind Machine Project
Ms. Rebecca Fahlstrom – Brain Sciences Foundation
Mr. Adam Howard – Brain Sciences Foundation
Mrs. Maryanne Jensen – Brain Sciences Foundation
Ms. Katie Johnson – Brain Sciences Foundation
Mr. Mazen Kishtban – Make a Mind Co.
Dr. Deepti Rao – Brain Coprocessors Inc.
Ms. Anastasia Semash – Brain Sciences Foundation

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form without the prior written permission of the Publishers.

Disclaimer
Statements of fact and opinion in the articles of The Brain Sciences Foundation are those of the respective authors and contributors and not of The Journal or The Brain Sciences Foundation. Neither The Brain Sciences Journal nor The Brain Sciences Foundation make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability for any errors or omissions that may be made.
# Table of Contents

## Visions

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain Language: The Fundamental Code Unit</td>
<td>Newton Howard</td>
<td>4–32</td>
</tr>
</tbody>
</table>

## Articles

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy Paradox of the Brain</td>
<td>Newton Howard</td>
<td>33–42</td>
</tr>
<tr>
<td>Cognitive Architecture: Integrating Situation Awareness and Intention Awareness</td>
<td>Newton Howard</td>
<td>43–59</td>
</tr>
<tr>
<td>Brain Space: Relating Neuroscience to Knowledge About Everyday Life</td>
<td>Newton Howard and Henry Lieberman</td>
<td>60–68</td>
</tr>
<tr>
<td>LXIO: The Mood Detection Robopsych</td>
<td>Newton Howard and Mathieu Guidere</td>
<td>69–75</td>
</tr>
</tbody>
</table>

## Analyses

<table>
<thead>
<tr>
<th>Title</th>
<th>Author(s)</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcranial Ultrasound Application Methods: Low-frequency Ultrasound as a Treatment for Brain Dysfunction</td>
<td>Newton Howard and Sergey Kanareykin</td>
<td>76–89</td>
</tr>
</tbody>
</table>

| Submission Invitation                                                                 |                           | 90      |
Editor’s Preface

It is our pleasure to introduce the first issue of The Brain Sciences Journal. This special introductory issue consists of six invited articles from Dr. Newton Howard, and collaborators, to showcase the journal’s objective in establishing a truly interdisciplinary neuroscience publication. In order to understand a collection of processes and functions as dynamic as in the human brain, our approach to study it must be equally dynamic.

This issue represents the most novel, forward-thinking research in the fields of brain sciences today. The six invited articles highlight a broad range of research that attempts to understand cognition, brain function, the root of intelligence, and neurological disorder. The first article, “Brain Language: The Fundamental Code Unit,” by Newton Howard, is a vision paper that sets the foundational premise for the rest of the issue. Brain Language is a paradigm for understanding and modeling brain function. It is perhaps the missing theoretical framework within the classical neuroscience approach, as it accounts for various interdisciplinary phenomena and draws on data from the last several centuries. “Energy Paradox of the Brain,” by Newton Howard, highlights the conundrum of the brain’s efficiency and capability. It is an account of a phenomenon observed by the author for several years, where in-silico models of the brain are far less efficient than that of their counterpart in-vivo.

“Cognitive Architecture: Integrating Situation Awareness and Intention Awareness,” by Newton Howard is a reporting on the last ten years of development in dynamic and static systems with the role of agency playing a significant part of system design. “BrainSpace: Relating Neuroscience to Knowledge about Everyday Life,” by Newton Howard and Henry Lieberman, is an attempt to benchmark and put to the test the most novel AI tools developed at the Mind Machine Project at MIT. The article recounts our understanding of commonsense computing and the broad interdisciplinary scope of how we can utilize these tools. “Transcranial Ultrasound Application Methods: Low-frequency Ultrasound as a Treatment for Brain Dysfunction,” by Newton Howard and Sergey Kanareykin, is a substantial review on one of the most relevant potential, non-intrusive methodology for brain stimulation and disease perturbation. “LXIO: The Mood Detection Robopsych,” by Newton Howard and Mathieu Guidere, is a demonstration of an architectural framework that integrates high level understanding of the inner workings of FCU, materialized with the implementation of the Mood State Indicator algorithm along with the role of agency. These six articles set the precedent for a new era in brain science, one that is driven by interdisciplinary inquiry and basic human curiosity.

Richard Wirt
Editor-in-Chief
Abstract

In this paper, we theorize that a Fundamental Code Unit (FCU) permeates the underlying workings of both the mind and brain's linguistic and communication systems, respectively. In this way, both the mind and the brain maintain schematics and axiomatic coding substrates similar to those found in DNA. Reconciling the paradox presented by the mind/brain duality concept promises to provide insight into several important phenomena, such as the location and composition of memories, as well as how they fit into the structure of the conscious mind, which offers a novel method of modeling human cognition that seeks low-level, indivisible components with which to construct human thought. In particular, the Fundamental Code Unit (FCU) is a blueprint for the construction of conscious thought, much in the same way that DNA is a blueprint for the proteins that drive biological processes. The FCU corresponds not only to the linguistic and behavioral output that results from cognition, but bridges the analytical gap between them and physical processes within the brain. In order to properly apply FCU to cognition, this paper argues that it is best to view the brain as a linguistic computational engine, using a topological rather than binary process for information analysis. It further proposes a unitary system, applicable for mathematical analysis on multiple levels. We use these tools to map out the process of cognition at each of these levels. This new approach to modeling brain function promises new insights into the apparent paradox of the brain's ability to consume low amounts of energy while generating high cognitive output. The concept of mind-brain duality argues that the mind is fundamentally different from the brain in the way it functions and operates.

Keywords: FCU, cognition, axiology, unitary, LXIO, deciphering, code theory, the DNA of cognition

1 Introduction

In order to offer a novel approach to formulating human cognition, important questions regarding the nature of human thought must first be addressed: what is the best way to define thought operationally, and what can this definition tell us about contemporary approaches to cognition? In the most fundamental sense, conscious thought consists of the transfer of signals between neurons. Since this description is not unique to cognition, which has basic processes that are neither chemically nor physically unique, any attempt to mathematically or physically comprehend human thought must be bounded accordingly.

Those who seek to understand the phenomenon of cognition must bear in mind the role of natural randomness in our observations of the human mind. In particular, the occasionally unpredictable nature of human thought does not lend itself to recreation by banks of computers run on software that accounts primarily for the performance characteristics of the human brain – that is, how large it is in bits, or how many different tasks it can perform at once.
For this reason, it is important to discern qualitative characteristics of cognition with which we can construct our understanding of it, and not just the observed capabilities of the human brain.

Our analysis of cognition seeks to determine not only its structure, but the units that comprise it. While many approaches to the study of consciousness privilege the connections between brain regions and cells as the primary vehicles of cognition, there still remains the question of how these connections are configured. We use the unitary system, or an intuitive series of unary mathematical functions, as the basis for modeling human thought processes. The applicability and simplicity of the unitary system suggests that cognition in itself is a monolithic process; that is, it is based on the same fundamental logic even though it is expressed in a variety of forms.

Drawing upon the unitary system and its biochemical correspondents, we propose the Fundamental Code Unit (FCU) as an important goal of our research. As DNA is commonly known as the “blueprint for life,” we argue that the FCU serves this purpose for thought; in researching its composition, we seek to show that like the proteins coded by DNA, cognition results from complex combinations of just a few components.

We argue that it is best to view the brain as a linguistic computational engine rather than a signal-processing engine since at the unit level, abstract cognitive concepts and their linguistic expressions reflect the same type of conscious thought. This is not to say that only humans who are capable of coherent speech process thought in this way (since this would imply that the cognitive processes of infants are somehow fundamentally different). Instead, linguistic expression is a reflection of brain function that can offer significant insight to researchers.

In addition, we view the problem of modeling cognition as a topological rather than binary process. That is, due to the fact that humans incorporate volition into their cognitive processes, a binary approach does not sufficiently explain the interactivity and interdependency of their thought processes. By organizing abstract concepts into linguistic expressions, humans have created numerous cognitive grammars that can be divided into atomic concepts and mapped to specific physical phenomena within the brain. Thus, examining the problem as one of “signal processing” within the brain means ignoring many of the observable cognitive structures that would be useful in a working model of conscious brain processes.

We view the problem as such: in analyzing Morse code, humans analyze the input dots and dashes at multiple levels; that is, they form letters, then words, then full thoughts, and finally, a response to the input. Similarly, we view the problem of mapping physical brain activity to cognition as taking place across multiple levels of analysis, each with its own fundamental units. Instead of tying our hypothesis to specific biochemical phenomena such as the firing of neurons and neurological signaling, our approach seeks to coherently unite the multiple levels of analysis involved in human thought.

The exercise of reducing cognition to a fundamental unit that comprises a series of larger structures is important for a number of reasons. Because cognition, as we know it, is a human construct (i.e., a series of biological processes deemed to compose intelligence), our conceptualization of it should reflect the configuration of the human brain and the processes that contribute to thought. At the submolecular level, DNA replication and protein coding does not appear fundamentally different from other chemical processes; the atomic bonds and interactions involved adhere to the same physical laws as any other chemical interaction. However, the A-T-C-G molecular composition of DNA is no less important as a result; this is because DNA amounts to the arrangement of ordinary elements into units that comprise sophisticated organisms. Similarly, the FCU takes ordinary physical phenomena and arranges them into units that not only explain, but also physically compose thought.

2 What is Cognition?

The emergence of consciousness from evolutionary, developmental and functional perspectives has
been applied in many areas of study. However, an operational definition must still establish clear parameters and fundamental mechanisms of conscious cognition. At the outset it appears that contemporary research has gravitated towards the quantum and electromagnetic theories of consciousness due to the fact that so little is known about this phenomenon, as well as the fact that neurons, the 'building blocks' of the brain, simply do not possess the inherent properties to fill the same role for cognition. This is because to deem the neuron a basic unit of cognition would be a gross oversimplification. Regardless of their cognitive abilities, many organisms share this biological structure in common with humans. Human neurons themselves aren't qualitatively distinct enough to account for our intellectual abilities versus other organisms, so the difference must lie in a combination of the gross quantitative differences (humans have hundreds of billions of neurons, so the possible neural network combinations are greater than organisms with physically smaller brains).

Wang (2003) describes a model of human thought processing that views information and knowledge storage in the brain as a series of relations, or connections, between objects and their attributes. This approach is differentiated from the container metaphor approach common in neuropsychology and computer science, which posits that individual neurons are the containers for knowledge within the brain. This approach, known as OAR (Object-Attribute-Relation), suggests that it is the properties of connections between individual neurons that determine the potential for information retention in the human brain [1]. The paper is offered as a reference model for the computing/IT industry, as well as information scientists; there is a heavy emphasis on the comparison between human and silicon computing volume/capability. Implicit in these comparisons is the notion that the human brain performs essentially the same function as modern computing hardware, just at a radically different speed and magnitude, and with a greater breadth of tasks.

The relation between ontology and knowledge representation is explored in greater depth in Guarino's 1995 paper. As a starting point, Guarino treats the distinction between reasoning and representation as the same as the philosophical distinction between epistemology, or the philosophical study of the limits and scope of knowledge, and ontology, or the study of nature and reality as a whole. In this sense, epistemology is ontology viewed through the lens of human perception, just as human representation through expression is equivalent to reasoning viewed through the human cognitive filter [2]. Traditionally, scholars of artificial intelligence and engineering tended to ascribe knowledge to agents or actors who could be observed pursuing a specific task with some sort of coherent strategy. However, Guarino argues that domain analysis ought to be considered a task-independent activity because there is often a multitude of task types and solution types, in addition to the occasional necessity for communication and collaboration between agents [Ibid]. Thus, to have such a rigid definition for knowledge or for intelligence would necessarily exclude significant amounts of coherent knowledge-based activity. Human memory is anchored in the concept of the locally observing and locally acting automaton; thus, it is not unexpected to see models of intelligence that proceed from this perspective. However, it is not clear that models for human memory would be the best place to start if one wishes to engineer memory appropriate for observing and actuating in a complex and distributed setting.

Gero (1990) proposes a multifaceted model of knowledge in his approach to interface representation design. In particular, his conception of relational knowledge is similar to that of Wang. Here he argues that variables such as function, structure, and behavior, while important in themselves, are largely interdependent, and it is the relationships of interdependency that form the bulk of useful relational knowledge. Qualitative knowledge provides information on the hypothetical alteration of the interdependent
variables of function, structure and behavior [3]. In particular, this type of knowledge informs the decision-making process by calling upon prior experience to compensate for uncertainty. Computational knowledge symbolizes mathematical relationships between key variables, and constraints include knowledge of certain actions that will affect the range of possible outcomes. Finally, context knowledge relates to external variables (characteristics of the environment, etc. that could affect the outcome in question) [Ibid]. What is particularly significant about Gero's approach is his perceived need to assign different qualities to different types of knowledge; it is as if at a unit level, these types he describes have little in common, and all are needed for coherent cognition to take place. To this end, one might view each ‘type’ of knowledge he describes as a basic unit of thought, as A, T, C, and G are fundamentally distinct but interdependent units of the DNA molecule.

Block (1962) describes the “perceptron,” or a series of sensory and associator units connected to resemble sensory and analytical components into a machine that vaguely models human response to sensory stimuli. Stimuli of a certain threshold trigger activity in specific associator units, which then activate those to which they are directly connected. Thus, different types of stimuli activate different networks of associator components. In this sense, Block's perceptron approach to modeling brain function is similar to our own in that it privileges the connections between components rather than the components themselves as the primarily important in decoding human thought [4]. However, there still remains the question of what constitutes a basic unit of connectivity. Does a single connection between two associates constitute a fundamental unit of perceptron “thought?” Studying the structure and function of different types of neural connections promises significant contributions, but this still begs the question of whether these connections consist, at some level, of units that are part of the same small set.

Anderson et al. (1998) developed the ACT (adaptive control of thought) a computational/mathematical model of cognition. This theory has multiple modules: the perceptual-motor modules, the goal module, and the declarative memory module. Like the perceptron, this model views cognition as consisting of several distinct subtypes that are not necessarily composed of the same units [5].

De Waal et al. (2010) perform research that supports the idea of cognitive continuity between human and non-human primates. According to them, mental capacities are based on a ‘deep homology’ of ‘the basic building blocks of cognition’ (such as homologous brain regions in birds and mammals) [6]. Instead of asking which species can perform at a specific cognitive level, they suggest that a better question to ask is what is necessary for that level of cognition to take place and how does it work? They suggest that mental capacities have to be broken down to neural level. They use mirror neurons as a bottom up approach to explain fundamental units for motor tasks, memory, future planning, pro-social behavior and imitation. This theory is more biological as opposed to the mathematical theory that comprises the Unitary System and the FCU, since it accepts neurons rather than arrangements of cognitive structures as a fundamental unit of thought. However, the two approaches are similar in that they both proceed upward after defining a fundamental unit.

Aerts (2009) discusses the cognition of concept formation and communication of meaning. They show that it is governed by the presence of quantum structure in cognition. They show that they can derive structure of human thought by interpreting quantum representation in a model that centers on “emergence due to superposition,” “interference” and “specific quantum field theoretic aspects [7].” According to them, human thought comprises two layers, the “classical logical layer” and the “quantum conceptual layer.” This model is not as concerned with pure linguistic structure. Rather it focuses on the “meaning aspects” of concepts and their combinations. This has significant consequences for the traditional problem of artificial intelligence, and the idea that meaning must be ascribed to the computational prerequisites of intelligence.

Lamb (2010) introduces the concept of the functional web, in which he posits that cognitive concepts such as single words and ideas (analogous
to semantic primitives) are in fact spatially distributed across parts of the brain such as the cerebral cortex. Lamb splits these concepts into conceptual, motor, phonological image, tactile, and visual components, or components that roughly align with the senses [8]. This approach not only applies to cognition but also to the concepts that comprise it is intuitive since its criteria are empirically grounded. In addition, it unifies behavioral and linguistic with neurological activity. What the FCU seeks to do in addition is to tie cognition not just to specific sensory activity but also to brain activity in itself. Lamb’s approach is more focused on response and activation, but the nature of cognition is such that thought can beget more thought; an external agent is not consistently necessary.

Blais et al. (2011) argue that modeling cognitive activity based on synaptic modification depends in large part on how synapses are stabilized after firing. With respect to synaptic activity, there are numerous types of “learning,” each of which has a different neuronal effect. Hebbian learning, for instance, occurs when the connectivity between two neurons decreases after one produces an action potential in the other. It is represented in a matrix as such [9]:

\[
\frac{dm_i}{dt} = \phi(c)d_i
\]

The selectivity-learning rule, on the other hand, incorporates a variable threshold of activation because it modulates the type and level of response to sensory stimuli (for instance, the difference between looking at the sun or at the night sky)

\[
\frac{dm}{dt} = c_k(c_k - \theta_m)d_k
\]

Blais et al. demonstrate an important mathematical connection between biology and temporality, or the idea that modeling such processes as cognition involves the accounting for change rather than for absolute physical values [9]. In this way, Blais does for synaptic selectivity what we do with cognition as a whole, and in doing so demonstrates the process parallelism that pervades natural phenomena. The unitary logic-based FCU is a more generalized form of Blais’s analytical approach; that is, it does not exclusively deal with stimuli.

Penrose (1991) suggests that consciousness is not just complex computation, but instead quantum computing of objective reduction, a self-collapse of the quantum wave function as a result of quantum gravity. He links the mind to space-time geometry. However, Penrose never specified the anatomical location of the seat of consciousness in the brain [10]. Hameroff et al. (1996) suggest that microtubules in neurons are the seat of consciousness and perform quantum computing [11]. Microtubules have proteins that associate with their quantum conformation states and thus orchestrate their actions. Thus Hameroff et al. calls the quantum states “orchestrated objective reductions [Ibid].” According to Hameroff et al., microtubule units can undergo quantum coherence across large areas of the brain. The quantum states of microtubules are maintained without environmental entanglement for up to 500ms. Microtubule units are able to interact with neighboring units. Conformational changes in microtubules can regulate synapses and other neuronal activities. This mode is related to unconscious activity. Quantum coherence among microtubule units occurs through electrons in hydrophobic pockets acting as a quantum computer. The associated proteins communicate the quantum state to the environment and hence collapse microtubules, resulting in what is known as orchestrated objective reduction. The selection of quantum states is random. Orchestrated objective reduction binds different microtubule unit conformations over large areas and at different time scales into a conscious though.

In a general sense, orchestrated objective reduction is a form of the popular collapse of the wave function theory that asserts that the quantum state activity in the microtubules is sufficient to influence synaptic transmission; thus, it represents a form of active conscious control of mental function [11].
This idea has been criticized because the calculated quantum decoherence happens within a time frame that would not be biologically significant. Other theories have responded by bolstering support for orchestrated organic reduction by factoring in quantum correction.

Using orchestrated objective reduction as a template, we still require results from a testable theorem. For instance, there are many clues that indicate that conscious processes are necessarily supported by a distributed network of synchronous neuronal firings. According to the endogenous electromagnetic field (EMF), conscious electromagnetic information field (CEMI) theory, and quantum brain dynamics (QBD) theory, among others, the synchrony of the action potential firing creates a local electrical field that is the substratum for consciousness. Cortical gamma oscillation, as well as the observed effects of deep brain stimulation, TMS and other electromagnetic manipulations support the notion that the brain does produce an EM field that can interact with other fields that influence brain function indirectly. While each of these theories contributes to our understanding of the complexity of consciousness, none of them can explain the phenomenological experience of agency, or the qualitative distinction that marks the conceptual gap between the human brain and the human mind. Intention Awareness is a subfield that approaches this problem more directly [12] and with recent research has led us to believe that, based on quantum spin-theories of consciousness, we may need to search for an unprecedented form of matter interaction due to the dearth of theories that can be used to explain cognition. Prior to such a search, however, the problem of cognition must be approached more directly.

Similarities between the Unitary/FCU approach and quantum coherence lies in consciousness governed by quantum effects. Hameroff et al. describes microtubules in a minimum of states of conformation that trigger consciousness. However, the advantage the unitary system has over Hameroff et al. in this is that it works at both quantum and cognitive level and explains quantum information propagation. Also, Hameroff et al. states that quantum collapse occurs on a scale of 500ms, which is not biologically possible, whereas cognition modeled under the unitary system is defined by different time scales [13]. In addition, the unitary system is not spatially limited or level-of-cognition limited as Hameroff’s microtubules are.

Contemporary conceptualizations of cognition often focus on cognitive structures and its overall character, but examinations of its physical nature at the unit level are equally important. Baslow (2009) frames his argument as a study of how receptors receive and transmit information. This is because, as he argues, there are two leading schools of thought regarding neuronal communication that are both difficult to verify empirically. The first, “spike timing code” theory, suggests that specific information is carried not only in the average firing rate of a neuron, but also in the precise timings between each spike. Alternatively, the “spike rate code” theory, posits, “information is coded by the mean firing rate of the neuron, and that spike timing is naturally random. Instead of engaging critically with these theories directly, Baslow explores the implications of both of them for the theory that language persists even at the interneuron level. Baslow calls the sending language a “two letter alphabet consisting of spikes and pauses between spikes that form specific meaningful reproducible “words” that can be used to transmit information obtained from a variety of external and internal sensors” [14]. The main advantage of applying language at this level, Baslow argues, is its universality:

“The universal nature of neuronal S/P languages is clear in that they can be written in many different frequency-encoded formats, and can be translated into any number of spoken or symbolic human...
languages, all with the same meaning. Neurons are also highly specialized communicating cells that use both electrophysiology, and what appears to be a more primitive chemical communication mechanism between neurons at chemical synapses. Thus, it follows that inter-cellular chemical communication that is used by a variety of modern cells as well as neurons is an ancient trait, and that all cells are probably biosemiotic systems” [14].

Baslow further develops the neuron as a semiotic object, or one capable of transmitting meaning, based on spike action potentials. Each neuron uses a simple signaling mechanism for communicating with neighbors: that electrophysiological action potential spike, which lasts about 1 millisecond, along with pauses. “Phrases” are repeated groups of spikes and pauses that compose words. Based on the S/P language he develops for interneuron communication, Baslow concludes that the basic “unit” of cognition in the brain is a dyad of connected neurons and their dendrites and synapses [14]. Two important pieces of evidence underlie Baslow’s claim. First, regardless of how information is communicated across synapses (rate or timing), it is clear that the spikes and pauses are its medium. Second, since a dyad of neurons is the smallest group across which the S/P language can be used, it must be by definition the fundamental unit of cognition.

For Baslow, there are four potential means by which neuron dyads can modulate signals sent: the length or thickness of the dendrites, and the length or width of the synaptic cleft. However, any change in the dendritic field can bring about a change in meaning of the words and phrases transmitted as spikes and pulses. Thus, Baslow argues that word-based translation takes place at the neuron level, and that words that are transmitted from one neuron to another are physically translated into a new spike/pulse language before they are interpreted by the receiving neuron.

This activity of translation allows for the recall of memory engrams encoding specific sensory recollections; this is because the act of recalling and the memory recollected will occur in different spike/pulse languages, but the meaning of the former will be preserved in the retrieval of the latter. In addition, Baslow asserts that anomalies in this translation process are the mechanisms of pathological and drug-induced changes in brain activity.

The binary basis for interneuron communication Baslow puts forth lends credence to the unitary system as a useful formula for decoding cognitive activity mathematically. However, his application of a broad definition of language (one that encapsulates both signaling between neurons and speech between humans) leads to a conclusion that is divergent with our own; namely, that the basic unit of thought must be defined in terms of biological structures that compose the brain.

Many of these approaches to cognition appear to assert the need for a conceptual breakdown of cognition in order to better model it. That is, each expression, whether linguistic or otherwise, reflects some contextual information, prior knowledge, and a rational schema based on qualitative and quantitative analysis. Since we posit the notion that each type of knowledge or form of thought could originate from the same fundamental cognitive units (or, to use a computer science analogy, data structures), it is important to address these positions. We do not believe that we are inherently incompatible with them; rather, their models simply don’t have a low enough level of analysis to acknowledge the possibility that the diverse types and numbers of knowledge and processing they propose are, at a basic level, woven from the same cognitive fabric.

To understand how the brain achieves these objectives on a consistent basis requires investigation of the problem at several levels simultaneously: from electrophysiology to molecular mechanisms of information storage and signaling, to neural network formation and trends. From a code and process to automata of sensory information, from the functional processes of perception, recognition, and behavioral control that must be implemented by the brain to the language expressed by higher intelligence and adaptation.
3 The Role of Linguistics in Deciphering Cognition

We posit that the linguistic base that humans possess is shared and rooted in the same natural language regardless of the language each of us speaks. While some of any natural language has evolved over time to include sophisticated semantics, there is a series of core grammatical and lexical concepts that have remained the same over time. These are linguistic ‘primes,’ which exist both at the word and sub-word level. In the latter case, there is a demonstrable analytical bias towards specific letter and phoneme types among francophone and other western language speakers [15]. Specifically, one study has found that “french and italian adults were found to be able to track transitional probabilities at the lexical level in a context of fixed consonants and variable vowels, but not the other way around” [15]. New et al.. (2008) concluded, “the scope of the consonantal bias at the lexical level is not even limited to the speech modality. It actually extends to lexical access through reading.” While this does not suggest that cognition is coded at the phoneme level, it does lend credence to the notion that, like biochemical reactions, cognition is composed of a few basic building blocks that can be arranged into any number of small concepts or words, which can in turn be combined into any of a rich set of thought expressions.

At the grammatical level are embedded semantic primes, a concept similar to the prime in the natural numbers (2, 3, 5, 7, 11, and so forth). Semantic primes, like prime numbers, are unique and cannot be constructed with other concepts; in this sense, they are truly atomic. By studying the low-level, indivisible conceptual constructs of languages, we can better understand these thoughts and perceptions. In addition, specifying a cognitive prime that plays a similar role in thought that the semantic prime does in language will better elucidate the connection between the two.

Semantic primitives, or innate concepts developed in the human mind and encoded in spoken language without explicit definition, provide a very basic interface between the human mind and its physical surroundings upon which cognition builds throughout one’s lifetime. They are an intuitive point of departure for analyzing the links between thought and language. Consequently, linking them to other cognitive concepts such as conception, perception and intention, it is possible to give semantic primes greater context and understand them better. By “marrying” data on semantic primitives to time-based, intrinsic, consequent and contextual value analysis, we argue that it is possible to better understand human cognitive processes and causes of behavior. The existence of semantic primes suggests that we have as part of our inherited human faculties a basic set of innate ‘concepts’ like intention awareness, or perhaps more precisely, a non-conscious propensity and eagerness to acquire those concepts and encode them in sound-forms (words) intentions. The words that those concepts become encoded in are called semantic primes, or alternatively, semantic primitives — ‘semantic’ because linguists have assigned that word in reference to the meaning of words. Words that qualify as semantic primes need no definition in terms of other words. In that sense, they remain indefinable. We know their meaning without having to define them. They allow us to construct other words defined by them. Anna Wierzbicka’s book, Semantics: Primes and Universals (1996), grounds this same argument in biologically plausible hypotheses and experimental observations [16].

As modern human beings who share the same set of types of inherited determinants that make us human (i.e., the human genome), all our natural languages, despite the diversity of language families, share the same basic set of innate concepts, or share the same propensity and eagerness to encode the same set of concepts in words. Our common grandmother and grandfather many generations removed may have encoded those concepts in a specific vocabulary, and therefore had an original set of semantic primes. The dispersal of their descendants throughout the world enabled the evolution of many different languages, each with a unique set of sound-forms for their words. Nevertheless the same set of semantic primes
remained within each language, though expressed in differing sound-forms. Thus, all modern humans have the same set of semantic primes, though not the same set of sound-forms (words) expressing them, rendering semantic primes cross-culturally universal [16].

4 The Biophysics of Cognition

The living neuron transmits information by “firing” it, delivering only an action potential. The inside of a resting neuron has a negative charge with respect to the outside. Sodium ions are actively pumped out of the neuron and potassium ions are pumped in. Potassium ions flow slowly across the membrane of the neuron, but sodium ions rarely cross it while the membrane is at rest. When the membrane is at rest, the electrical gradient and concentration gradient act in competing directions for potassium, almost balancing out. Both gradients tend to push sodium out of the cell. When the charge across the membrane is reduced, sodium ions can flow more freely across the membrane. When the membrane potential reaches the threshold of the neuron, sodium ions enter explosively, suddenly reducing and reversing the charge across the membrane. This event results in the action potential. For any stimulus greater than the threshold, the amplitude and velocity of the action potential are independent of the size of the stimulus that initiated it [17].

After the peak of the action potential, the membrane returns to its original level of polarization because of the outflow of potassium ions. Depolarization of the membrane opens voltage-gated sodium and potassium channels and the sodium channels snap shut at the peak of the action potential [17].

A neuron adds and subtracts excitatory and inhibitory inputs until it reaches a threshold, at which point it fires a single impulse or action potential. At the neuromuscular junction, virtually every action potential in the presynaptic motor neuron triggers an end plate potential in the postsynaptic muscle cell. However, the situation at synapses between neurons is more complex because the postsynaptic neuron commonly receives signal from many presynaptic neurons. A single neuron can be affected simultaneously by signals received at multiple excitatory and inhibitory synapses. Just as the brain constantly integrates the unary “plus” and “minus” of words when generating linguistic expression, so too does the neuron continuously integrate these signals and determine whether or not to generate an action potential based on the unitary system [18].

Generally speaking, atoms such as hydrogen, nitrogen, and oxygen, each have a unique marker that defines its behavior and its interactions with other atoms. This arises from protons, neutron, and electron counts and configuration and, more importantly, the ionization energy as defined by an atom's valence [19]. The ionization that takes place at the valence shell yields energy. If two atoms react and approach one another, it is their outer shells that become involved in any chemical reactions, such as the synaptic exchange. Thus, the ionization energy related to synaptic exchange is driven by the interactions amongst neuronal molecules.

Shifting concentrations of elements within the brain, particularly the activity of the ions, create an ebb and flow of charge and energy between individual neurons. The flow of electron and protons as action potentials throughout the brain matter gives rise to an energetic field that we call consciousness, which manifests in language as we have previously seen [20]. Thus, the Unitary System is applicable in modeling the changes in this field, which result from the output of chemical transactions between neurons. In particular, neuronal action potentials that drive synaptic communication are themselves driven by a rush of sodium ions into the receiving cell, a process that has a depolarizing effect on the membrane potential. Recovery, on the other hand, is affected by a rush of potassium ions through potassium
channels. Both of these affect what is known as the membrane potential of the receiving neuron. Membrane potential in a cell derives from two factors: electrical force and diffusion. Similarly, these are the two primary mechanisms of neuron communication. In the former case, the net actions of ion channels and ion pumps (such as the sodium–potassium ATP-ase) embedded in the membrane produce differences (and therefore electrical charge) on the intracellular and extracellular sides of the membrane.

From the viewpoint of biophysics, the resting membrane potential is merely the membrane potential that results from the membrane permeabilities that predominate when the cell is resting. At any given moment, there are two factors for an ion that determine how much influence that ion will have over the membrane potential of a cell. The first is the ion's driving force, which is the net electrical force available to move that ion across the membrane. It is calculated as the difference between the voltage that the ion “wants” to be at (its equilibrium potential) and the actual membrane potential $E_m$. So formally, the driving force for an ion is equivalent to $E_m - E_{ion}$. For example, if there is a resting potential of $-73 \text{ mV}$, the driving force on potassium is $7 \text{ mV}$: $(-73 \text{ mV}) - (-80 \text{ mV}) = 7 \text{ mV}$. The driving force on sodium would be $(-73 \text{ mV}) - (60 \text{ mV}) = -133 \text{ mV}$. Second is the ion's permeability, which is simply a measure of how easily an ion can cross the membrane. It is normally measured as the (electrical) conductance and the unit, siemens, corresponds to 1 C·s−1·V−1 that is one charge per second per volt of potential. If the driving force is high, then the ion is diffusing in one direction faster than the other. If the permeability is high, it will be easier for the ion to diffuse across the membrane. Thus, in a resting membrane, while the driving force for potassium is low, its permeability is very high. Sodium has a high driving force, but almost no resting permeability. In this case, potassium carries about 20 times more current than sodium, and thus has 20 times more influence over $E_m$ than does sodium.

The effect of increasing the permeability for a particular type of ion is to shift the membrane potential toward the reversal potential for that ion. Thus, opening sodium channels pulls the membrane potential toward the sodium reversal potential, usually around $+100 \text{ mV}$. Opening potassium channels pulls the membrane potential toward about $-90 \text{ mV}$; opening chloride channels pulls it toward about $-70 \text{ mV}$. Because $-90$ to $+100 \text{ mV}$ is the full operating range of membrane potential, the effect is that sodium channels always pull the membrane potential up, potassium channels pull it down, and chloride channels pull it toward the resting potential.

At the peak of the action potential, on the other hand, permeability to Na is high and K permeability is relatively low. Thus the membrane moves to near $E(\text{Na})$ and far from $E(\text{K})$. From this example, it becomes clear that the more ions are permeable, the more complicated it becomes to predict the membrane potential. However, this can be done using the Goldman-Hodgkin-Katz equation or the weighted means equation. By simply using the concentration gradients and the permeabilities of the ions at any instant in time, one can determine the membrane potential at that moment. What the Goldman-Hodgkin-Katz equations tell us is that at any time, the value of the membrane potential will be a weighted average of the equilibrium potentials of all permeating ions. The “weighting” is the ions relative permeability across the membrane.

In neurons, the factors that influence the membrane potential are diverse. They include numerous types of ion channels, some that are chemically gated and some that are voltage-gated. Because voltage-gated ion channels are controlled by the membrane potential, while the membrane potential itself is partly controlled by these same ion channels, feedback loops arise which allow for complex temporal dynamics, including oscillations and regenerative events such as action potentials. Each of these variables can modulate the polarity that results from each chemical transaction and thus drives synaptic communication.
The Brain Sciences Journal

The model for two populations of neurons to address connections to neighboring networks of actions, or neuron column, assumes a constant production efficiency, meaning potentiation or activation ratios. This efficiency also holds at \((-10\text{mv})\) with \((+\)\) displacement and a \((-20\) with \((-\)\) displacement as in the classical Lotka–Volterra equations, which succinctly utilize the energy flow principle: in essence, we are energy processors.

Here the system is divided into two levels, \((x\text{ and } y\)) are their respective densities or biomasses in the same units) and the yield constant or the production efficiency \((e\)), manifests the second law of thermodynamics \((0 \leq -e < 1)\). Since \(e\) is a constant and the functional response of \(1a\), \(f(x)\), is a monotone non-decreasing function, it follows that higher density never decreases the flow of energy to the next level. In fact, all relationships are considered as a \((+, -)\) type, as indicated by the signs of the off-diagonal terms in the community matrix or Jacobian of system.

At a synapse, a neuron releases neurotransmitters that excite or inhibit another cell or alter its response to other input. Excitatory neurotransmitters, the most common type, increase firing rate. An inhibitory neurotransmitter decreases the chances of the neuron firing. Each neuron is influenced via multiple neurotransmitters acting at multiple synapses by dozens of other neurons [21].

Following the release of a neurotransmitter and the subsequent activation of a receptor, it is important that the response is terminated and the system reset so that a subsequent activation can occur. This is achieved through the removal of the neurotransmitter by metabolic enzyme activity and by passive or active uptake activities. The concentration of the transmitter at the synapse for a longer time period occurs if the uptake mechanism is blocked. Therefore, a neurotransmitter uptake blocker may have an effect similar to a postsynaptic agonist of that transmitter. For uptake to take place, the neurotransmitter must be recognized by an uptake mechanism. As a result, it is common for structural analogs of the neurotransmitter of this process, for example, noradrenaline, serotonin, and dopamine. Here, we see the unitary system at work in the form of molecular chirality, or a molecular property deriving from a lack of internal symmetry.

Biological molecules such as neurotransmitters exist in mirror image isomers of one another, and this is what governs the ligand-substrate interaction specificity necessary for biochemical reactions [22][23]. The recognition of a neurotransmitter by its complementary receptor occurs due to the unique conformation of each isomer-enantiomer ligand, functioning as a “lock and key” mechanism. The two mirror image isomer-enantiomer molecules will interact with post-synaptic receptors, producing different biochemical effects. The ligand-substrate specificity required for neurotransmitter activity is conferred by the unique electronic interactions between asymmetric molecules [23]. The “all-or-nothing” action potential is generated by the summation of excitatory and inhibitory signals in the form of chiral molecules that are found in a \(S\ (+)\) or \(R\ (−)\) isomer-enantiomer conformation binding and releasing from the receptor, generating a strong enough signal to cause
the neuron to reach threshold. While the chiral neurotransmitters are mirror images of one another, it is important to remember that they are not the same molecule and will not exert the same strength of signal upon the receptor. The salience of the signal is determined by the strength of the ligand-substrate interaction [24]. Thus, the individual effects of each neurotransmitter are tied to its particular chiral pair configuration.

With regard to the central stimulant actions on a regional structure of the central nervous system, the S (+) isomer is several times more potent than its R (−) enantiomer. The S (+) isomer is known to induce euphoria, whereas enantiomer R (−) has been linked to depression [25]. The overall greater potency of the S (+) isomer form in such cases suggests that this form may have a higher potential for deep cranial stimulant actions and neurotransmitter availability in the synapse.

5 The Unitary System: bridging the gap between biophysics and linguistics

Parallelism in cognitive phenomena

Parallelism in physical processes is a cornerstone of the FCU argument. That is, basic building blocks of language (letters) are analogous to DNA’s role in protein formation, which is similar to the FCU’s role in formulating coherent thought. We argue that the underlying units that compose cognition, like those of DNA, are relatively simple compared with the structures they create. This applies both to the brain itself and the way we perceive it (i.e., as a system of sensory inputs and linguistic and behavioral outputs). In our approach, we map the physical phenomena of cognition to this theoretical system. The functions “unary plus” (+) and “unary minus” (−), representing an increase or decrease in the underlying measured value, carry sufficient computational efficiency to represent human cognition, provided that the same linguistic base is present on both sides. Thus, demonstrating a basic linguistic parity among human minds would alleviate the need for a number-theory based approach to the conscious brain function that we seek to model. Mathematically speaking, in between the semantic primes and unary constructs, there exists a set of functions, which together with the unary construct form a new system of mathematics: the unitary system. The brain communicates within itself and with the rest of the body via unitary operators. These unitary operators carry a state of time and space that conveys information necessary to decipher any semantic or non-semantic based language.

Since we postulate the existence of FCU and unitary system for communication within cognitive systems, we expect deeper examination to reveal correlations between that unitary system and the neurological functions at multiple levels, down to the smallest unit level of neurons. For example, the behavior of quarks, the smallest measurable particles known to man, is governed by 3 combinations of positive arrangements and 5 negative combinations [26]. When DNA actually replicates itself it does so on prime segments: 3 and 5, 5 and 3 [26]. It becomes apparent at the sub-atomic level that this 5/3 ratio relates to the quantum value of quarks residing within the sub-atomic particles, which in turn reside in the nucleus of the atom [27]. With the quark codes on the state of 3 and 5, and the level of efficiency in doing so allows the body to repeatedly produce DNA at a high daily rate. If we look at the
The Brain Sciences Journal

chemicals produced in the body, we see that a lot revolves around two arrangements of the atoms in the molecule – the enantiomer and the isomer; each chemical has 12 possible geometric configurations, with their combinations once again using primes 3 and 5 [28]. Once the specific chemical configuration is known, one can precisely engineer a medicine for anything that the body would need, either through stimulating the brain or correction of a damaged system.

The chemical action potentials that drive neuronal communication show the unitary system at work. This is due to the photoelectric exchange ratio, or the amount of photonic energy required ionizing an atom, or changing its charge state. For instance, if we assume that the ratio is 5:3, then 5 photons are required to ionize a given atom by 3 electrons. For example, if each photon has 1,759.2KJ/mol of energy, then those 5 photons would ionize iron into the +3 state. The 5 photons would have to hit the iron simultaneously though. We theorize, however, that each actionable photon should hold the quantum of energy required for 1 electron. Thus, using the above example, rather than having 5 photons at 1,759.2KJ/mol, we would ideally want 3 photons. Those 3 photons would have different energies. The first photon would have 759.3KJ/mol to exactly match the first ionization energy of iron. Photon number 2 would have 1,561KJ/mol to exactly match the second ionization energy of iron. And photon number three would have 2,957.3KJ/mol, exactly matching iron’s third ionization energy. Since ionization is a key mechanism of intercell communication between neurons, this ratio, modeled by the unitary system, can be correlated with specific cognitive activities.

A quantum logical gate can be described as a unitary operator, assuming arguments and values in a product-Hilbert space. From an intuitive point of view, it seems natural to see the gate as a kind of truth table that transforms + and −. The gate can also be represented as a matrix. One can observe the same quantum physics at work when they consider cognition at the meso-level by examining the physiology of the brain; this is particularly significant in that it demonstrates the multi-level applicability of the unitary system. That is, the unitary system’s underlying logic holds at both the quantum and cognitive levels.

Using the unitary system, information, coded using the FCU, travels from positive/negative cognitive constructs, to action potentials in neurons, to chiral molecular structures of neurotransmitters, DNA and proteins, governed in the end by quantum effects. On each of these different levels, the unitary system is being used to model the propagation information. Although it takes different forms at each transition, it can be described via a set of related mathematical formulas, rooted in number theory, which in turn exposes the most basic patterns and sequences of our universe. Understanding these patterns and transition will allow us to better understand the human mind, and to more precisely engineer specific compounds that the brain and body need to develop, heal, and perform better.

Information Processing

Does the brain use stochasticity in computation, or does it simply average out convergent neuronal input? Many scientists assume the latter, but those that challenge this assumption have to deal with a lack of mathematical tools. Active properties are based on the aggregate behavior of many voltage-gated ion channels that are in themselves stochastic mechanisms; tools already exist for the examination of these processes. The Shannon-Weaver Model of Communications (SW), for instance, states that there are three levels, or obstacles, that must be overcome when constructing a system of interaction [29]. In our case we are essentially reverse-engineering an existing system, but the claims made by this model are still equally relevant. In particular, this model calls for technical, semantic and effectiveness aspects of an interaction system to be resolved before it is deemed effectual [29].

The S-W model asks us how accurately the symbols of communication are transmitted; that is, if the rate of errors is low enough to maintain overall functionality. In the case of the human brain, it is clear that synaptic electrochemical mechanisms do a sufficient job of communicating signals between
brain cells, but does the Unitary System and FCU properly account for this? The S-W model also demands conveyance of the precise meaning of symbols, and for the message to affect behavior in the intended way.

With regard to the communication aspect of the S-W model, a lot of signal noise occurs in neuronal transmission; in fact, transmission probability (i.e., the probability of reception) can be as low as 10%, and thus the method of neuronal signal transfer in itself is not particularly reliable. This is compensated for by greater redundancy. More reliable and stronger synapses do exist, but short-term signal modulation is difficult with these, as if habituation to transmitting a specific signal or type of signal improves the synapse. On the other hand, short-term modulation capacity’s tradeoff is the fact that it depends on frequency and use-history.

In neurons, information is transmitted as “spikes;” it is unclear how exactly these spikes encode information, and there is a continuing debate in the field about whether the timing of these spikes is an important informational component. Conventional wisdom argues that the average firing rate is more important, and this is due to the apparent correlation between stimulus strength and firing rate. However, the observation of neuron firing rate by external measurement has some built-in latencies that make accurate measurement of these figures difficult [14].
Those who support the notion that neuron spike timing is important fall into one of two camps. First, there are those who argue that interspike intervals in the activity of a single neuron can potentially transmit information. In addition, others believe that the relationship between spike timings of different neurons may be important to the conduction of information. The latter argument is based primarily on the notion that synchronized neuron firing is an effective mechanism for activating postsynaptic cells, and is therefore an effective means for encoding information, especially in dynamic environments [Ibid].

It is very difficult to conduct experimentation on neuron spike timing because statistical approaches such as principal component analysis don't offer insight into the coding strategies of neurons. Consequently, there needs to be significant progress in statistics and theory of stochastic information processing before a conclusion can be reached.

Proceeding from our prior assertions that the brain is fundamentally a linguistic computational engine and that its processes can be modeled topologically, we argue that the qualitative and quantitative specification of a Fundamental Code Unit is a prerequisite to a robust model of human cognition. This is because it will allow more accurate computational and mathematical reproduction. The FCU is not specific to one level of cognitive analysis (i.e., biochemical, quantum, psychological etc.) but, once specified, serves as a building block for the construction of new analogous thought structures outside the brain.

In order to support the notion that the FCU is the building block of thought, we must seek specific arrangements of phenomena within the brain. Within the brain, the FCU will correspond to the biochemical and molecular interactions that govern human thought. That is, synapse interaction and activation, as well as the protein exchanges between neurons, will comprise the FCU. In what way they do so, however, will depend on how the structures of linguistic expressions and cognition are reflected at this biochemical level. Thus, we seek to intuitively connect the simplest, most fundamental components of thought with the sophisticated, intelligent cognition we experience as thinking humans. The goal is thus to specify both the quantitative characteristics (i.e., how many molecular interactions or semantic components are needed to generate thought?) and the qualitative characteristics of cognition (i.e., what do FCU components and processes do in concert that would not be useful to our understanding of cognition if they were considered alone?).

If cognition is composed of like units, each must account for the various sources of conscious thought. The act of thinking, or reasoning, calls upon events in both long and short-term memory, in addition to applicable learned concepts. Because language maps thoughts to a coherent expression, linguistics is a potent source of cognitive insight. The FCU takes this relationship between thought and linguistic expression a step further by mapping the underlying physical processes that ultimately lead to this expression of language and behavior. Since cognition is a spatial as well as temporal phenomenon, the units that describe and comprise it must reflect both of these dimensionalities. Thus, we introduce the mind state, or a physical approximation of the human mind or a subset of it at a given point in time, and use the Markov Decision Process model to map its transitions through reasoning and decision-making.

In cognitive processes, predictions based on the present are not made any more precise through additional information on the past. All useful information for predicting the future is contained in the current state of the process. One of the most useful tools for studying this property is the Markov chain. A Markov decision process (MDP) allows individuals to make decisions when there is inherent certainty about the current state and uncertainty about the effect of their actions. Assuming that a state of mind evolves over time, then any change has a given probability of being positive or negative. Now, should one want to control the state of mind, there is a range of
possible actions (e.g., therapy) that, in turn, are also probabilistic. The efficiency of the control exerted is measured based on the positivity or negativity of responses throughout the experiment. The process can be summarized as follows: when in state $s$, there is probability $p$ that, having chosen action $a$, the subject finds himself in [new] state $s'$ with benefit $b$. The progression of mind states corresponds to a Markov chain. In other words, the mind follows a sequence of distinct states over time, based on the probability of each transition. According to Markov’s hypothesis, transition probabilities depend exclusively on $n$ previous states. Generally, $n$ is equal to one in order to restrict analysis to the current state and the following state.

A Markov process is an $\{S, A, T, R\}$ quadruplet where:

- $S = \{S_0, \ldots, S|S|\}$ is the finite discrete set of possible states
- $A = \{a_0, \ldots, a|A|\}$ is the finite discrete set of possible actions
- $T: S \times A \times S \rightarrow [0; 1]$ is the function describing the change in state resulting from the control actions

Generally, the $T$ function is probabilistic: $p(s'|s,a) = T(s, a, s')$ being the probability that the mind goes from state $s$ to $s'$ when action $a$ is undertaken. $R: S \times A \times S \rightarrow \mathbb{A}$ is the reward function indicating the real product of choosing action $a$ while in state $s$ in order to achieve state $s'$.

Figure 5 represents a MDP for three different states $\{S_0, S_1, S_2\}$, shown in green. From each state there is the possibility of choosing an action from the set $\{a_0, a_1\}$. The red nodes therefore represent a potential decision (the choice of an action in a given state). The numbers indicated on the black arrows indicate the various transition probabilities for each decision node. Finally, the transitions can generate rewards, as shown by the upward yellow arrow (+5) and the downward yellow arrow (-1).

The transition matrix drawn from $a_0$ is:

$$
\begin{pmatrix}
0.50 & 0 & 0.50 \\
0.70 & 0.10 & 0.20 \\
0.40 & 0 & 0.60
\end{pmatrix}
$$

The transition matrix drawn from $a_1$ is:

$$
\begin{pmatrix}
0 & 0 & 1.0 \\
0 & 0.95 & 0.05 \\
0.30 & 0.30 & 0.40
\end{pmatrix}
$$

With regards to rewards, there is a +5 reward when going from $S_1$ to $S_0$ by doing $a_0$ and a -1 reward (also called penalty) when going from $S_2$ to $S_0$ by doing $a_1$. Using this mind-state mechanism, we argue that it is not only possible to approximate cognitive processes in a computationally efficient manner, but that these same processes occur in the physical realm.

Taking the analysis a step further in the spatial direction, we introduce an algebraic component. Begin with an infinite set $S$ representing brain regions that may be activated by some means. We introduce a $\sigma$-algebra $\mathcal{A}$ on that set, and call the elements $a \in \mathcal{A}$ activation sets (by definition, $a \subset S$).

Next, we introduce a second set $W$ whose elements are labeled concepts in the brain that correspond to words. For some subset $A \subset \mathcal{A}$ there is a mapping $P: a \in A \rightarrow w \in W$ called the concept activation mapping. The elements $a$ of $\mathcal{A}$ are action potentials. Let $P: \in W \rightarrow \hat{a} \in \hat{\mathcal{A}}$ be a mapping we call the brain activation mapping. Let $\mu$ be a measure on $S$ and let $F: \mathcal{A} \rightarrow \{+,-\}$ be a parity mapping. An axiology is a mapping $\Xi: W \rightarrow \{+,-\}$ generated by computing $f(w) = a \Xi F(s) d(\mu)$ with $a = P(w)$ and then projecting $\Xi(w) = \text{sign}(f)$.

![Figure 5: A MDP Showing Transitions Between Three Mind States.](image-url)
This system describes the mental mood states of an individual by examining the mind's abstract structures. Throughout the brain there are various forms of activations (electrical, chemical, biological) each contributes individually or within groups to the formation of new concepts, which define a positive or negative mental state. Series of such activations form an activation set, this set represents a connected structure for each activated region, which is defined in terms of a node. The node circumference changes based on the duration of the region being activated and the reflexivity is due to the re-activation of this region at different instances.

6 Maximum Entropy Statistical Analysis of the Unitary System and FCU

The Maximum-Entropy (Maxent) model of statistical analysis can be used to test the fidelity of the predictions made using the Unitary System, as well as to determine how much and what type of data should be encapsulated in the Fundamental Code Unit. Maxent is a discriminative, or conditional, model that accepts input data and uses its inherent patterns to infer about missing data - in this way, the results of a Maxent analysis become more accurate not only in proportion to the amount of data available, but the type of data as well.

To construct a Maxent model for Unitary System analysis, we must establish several prerequisite features, or pieces of evidence tying observations (i.e., empirical data), to our predictions about them. In Natural Language Processing (NLP), these features often include the specific word being analyzed, its shape (hyphens, suffixes and prefixes etc.) nearby words, and their features. [30] In the case of the FCU under the Unitary System, we can use Maxent not only to examine the features FCU at the unit level but also the correlation between Unitary System mathematics and the process by which fundamental code units are transmitted within the brain, and then expressed behaviorally. Thus, we are testing two separate analytical levels using Maxent. The first is the qualitative composition of the FCU, or the features it encompasses. Comparing the FCU to linguistic expressions and other behavioral observations will allow measurement of the conceptual preservation of the FCU. We are also measuring similar preservation between the Unitary System’s unary function-based model and the observed spike-pulse alphabet that we observe within the brain.

7 Application for Unitary System and FCU

The unitary system promises a straightforward approach to modeling cognition as an intuitive
Figure 7: The Translation of Physical Into Cognitive Phenomena
mathematical model. This system is applicable at multiple levels of brain function, from language acquisition to neuronal interaction to DNA replication and the subatomic processes that govern it. We argue that the strength of the unitary system lies not only in its wide applicability in the cognitive realm but in the new opportunities it offers to researchers and scholars of cognition: the ability to discern a fundamental unit of thought. Combining the unitary system with empirical data promises to help determine not only the quantitative minimum of cognition (i.e., the number of active neuronal connections that form conscious thought), but its qualitative foundations as well.

**Ultrasound**

Ultrasound is well established as a diagnostic imaging tool but its therapeutic potential, particularly for central nervous system disorders, has been largely overlooked until now. Recent technological advances now allow transmitting and focusing of ultrasound through the intact human skull using an array of phase-corrected ultrasonic transducers placed on the cranium [31]. As the technology matures and the control over brain effects becomes more fine-grained, an understanding of the structural relationship between cognition and its physical media will allow greater therapeutic strides to be made.

Such non-invasive, focused ultrasonic intervention permits thermal (high power) and non-thermal (low-power) modes. Non-invasive, thermal ablation of thalamic nuclei using ultrasound has recently been demonstrated to be effective in the treatment of neuropathic pain patients [32]. We argue in favor of the use of non-invasive, focused ultrasound in non-thermal stimulation and suppression of neural activity [33]. We hypothesize that non-invasive, transient and focused ultrasonic neuromodulation (USN) offers great potential for the therapy of psychiatric disorders that are associated with dysfunction in defined anatomical structures but evade effective pharmacological treatment.

**Discourse Analysis**

Within the realm of cognitive studies, we posit that one can analyze a state of mind as “positive” or “negative,” which can also be expressed along the time axis, as being future-oriented or past-oriented. Human cognitive architecture allows us to digest information as negative or positive values. However, it is by processing our cognitive structure with this cognitive architecture that we can judge whether a given concept has a positive or negative value. During discourse, we unconsciously select words, which signify positive or negative concepts [34]. Thus, unary functions, i.e., ‘+’ and ‘−’, allow efficient processing of language information provided that you have the same linguistic base on the communicating parties. The linguistic base is essentially the sum of any natural language, which evolved over time to include semantics with level of sophistication and potency, built on the foundation of embedded semantic primes. The semantic primes, i.e., concepts that are understood intrinsically, have positive or negative values in the shared value system, or axiology [34].

Analyzing discourse, one can deduce positive and negative values for specific words, based on the intrinsic value, time orientation, context, and frequency. Internal or inter-personal discourse, associating concepts with their positive or negative values (e.g., abstracted as reward/punishment, pleasure/pain, etc.) drives the actions and attitudes, and can thereby be quantified, analyzed, and corrected. This has applications for diagnostics and treatment of mental disorders, starting with depression, and general implications for cognitive modeling. Mathematically speaking, in between the semantic primes and unary constructs, there exists a set of functions, which together with the unary construct form the new system of mathematics – the unitary system. The brain communicates with other organs, and within itself, in unitary operators. The unitary operators carry a state of time and space, conveying all that is necessary to decipher any semantic- or non-semantic-based language, including pictorial languages.
Figure 3 depicts the underlying actions that take place in the brain at the various levels from biological, electrical, chemical through the cognitive and linguistic levels. Here, we are trying to trace how a word, in this particular example “Life” can be viewed by a patient and what formulates the mood state associated with it. Words are processed in the brain based on their letters and how each letter relates to the other from a value standpoint. When analyzing a word value we cannot disregard the fact that each group of words can represent a concept that could be significant to understanding a patient’s state of mind. Looking closely at what triggers this mechanism of word valuation and mood identification, we first examine the atoms that react and interact with one another while their chiral molecules produce different biochemical effects that in turn generate electrical signals. The signals characteristics specify the isomer-enantiomer molecules that result in euphoric or depressive state of mind respectively. These mood states are associated directly to the energy generated at various signal spectrums.

The resultant S+ and R- will trigger chemical secretions at the synaptic levels and uniformity of such interactions will result in a unary operation that compensates for the elevated levels of such chemicals. Only signal frequencies that are involved in this particular evaluation are considered. Hence, frequencies < 4Hz are only taken in consideration for the sake of our example. In other cases, we look at different frequencies associated with the neuron activation. Particularly, in some instances we might need to consider these signals at different time frames and frequency oscillations due to their direct affect in determining the value of each letter or word. The operation that takes place at the synaptic level generates various resultant signals, which when correlated generated signal amplitudes that represent various letters of the alphabet. This unitary system produces particular signals with amplitudes and oscillations that are directly associated with letter values and mood characterization. After each letter has been thoroughly examined and analyzed compared to the previous and preceding letter, it is subjected to a value assignment which when calculated as a whole produces a positive or negative axiological value that represents a mood state at that particular time instance within a specific context. Although the various words form concepts some, especially if they are universally accepted, form conceptual metaphors, which are evaluated using the same framework, described here.

A practical application of this form of discourse analysis is the LXIO cognitive mood analyzer, which parses sentences into words that are processed through modules to evaluate cognitive states. Words are processed through time orientation, contextual-prediction and consequent modules. LXIO relies on two core concepts in its formulation of mood: first, human states of mind are a spatiotemporal phenomenon. That is, the human mind’s progression from one state to the next can be quantified by variation in these units, which in turn forms a pattern of change over time that can be mapped to coherent thought.

Regardless of language, LXIO takes into account linguistic expressions. LXIO is a word-value-based analyzer that assigns positive and negative values to each parsed word based on the modules. The primary metrics for LXIO analysis include the time frame, orientation and concept set of the words spoken by the subject. This way, the activation of certain concept sets can be correlated with environmental and physiological variables in order to determine the causative agent in thought process, as well as its future progression.

As an alternative to the compartmentalized notion of cognition many researchers propose, LXIO is a testable means of insight into the relationship between language and underlying cognitive structures; it promises to yield techniques for automated diagnostics of PTSD, depression and similar symptoms, as proof of concept [34]. Using metrics such as the cultural or traditional significance of spoken words and phrases, the frequency with which they are used, and their context relative to other phrases, LXIO projects
the “value” analysis of a given statement. Thus, in its analysis of mood state and diagnostics, LXIO subjects all input to the same fundamental mathematical transformation, and in doing so it produces positive clinical results [34].

With respect to cognition, the significance of LXIO is its ability to incorporate conscious thought and its bodily expressions, linguistic or otherwise, into a uniform code schema. The ability of LXIO to diagnose brain conditions based on mood analyzed through speech lends credence to our view that the whole of cognition, including language, consists of a series of different structures arising from the same units.

*fMRI Neurofeedback*

Shibata et al. (2011) describes a fMRI neurofeedback method for inducing visual perceptual learning that bears relevance to our position in that their findings contain two important implications: First, visual perceptual learning (VPL) in the Early Visual Cortex of adult primates is sufficiently malleable that fMRI feedback can influence the acquisition of new information and skills when applied to the correct region of the brain [36]. Second, these methods can induce not only the acquisition of new skills and on formation but can aid in the recovery of neurological connections that have been damaged by accident or disease. For instance, a trauma victim suffering from language skill loss can potentially recover those skills through fMRI neurofeedback induction.

What Shibata’s paper shows us about the structure of thought is that the FCU, which we seek in cognition must be based on some finite number of neurological connections – those same connections influenced by the activity of fMRI neurofeedback. This process does not target a single neuron, but a locality of connected neurons, and based on its positive effects on the conscious process of Visual Perceptual Learning, the FCU represents that reality. In addition, Shibata’s fMRI induction research can provide powerful evidence for the composition of thought because it can be used to determine the minimum amount of neuronal connectivity for the formation of thoughts.

*Toward a superior computational understanding of brain function*

In order to study brain function, some researchers have attempted to reverse-engineer neuronal networks and even the brain itself. This approach was based on the assumption that neurons in-vivo acted just like simple transistors in-silico. Unfortunately, both network and whole-brain modeling have led to very little insight into actual brain function. This is largely because transistor-based computing reacts to static events whilst neurons can react to processes. In contrast to transistors, neurons can establish and change their connections and vary their signaling properties according to a variety of rules, allowing them to adapt to circumstances, self-assemble, auto-calibrate and store information by changing their properties according to experience. Consequently, a detailed understanding of neuronal function and network organization is required prior to neuronal network modeling attempts.

Economy and efficiency are guiding principles in physiology [37]. We thus argue that understanding the design rules which shaped neuronal function and network organization will help us in understanding and modeling of brain function. This notion was already recognized about 100 years ago by Ramon Y Cajal, who stated “all of the formations of the neuron and its various components are simply morphological adaptations governed by the laws of conservation for time, space and material” [37].

The structural and physiochemical relationships that connect resource use to performance in the brain are determined by three major constraints: geometrical limitations on packaging and wiring, energy consumption considerations, and energy-efficient neural codes.

One obvious difference between current in-silico technology and the brain is their three-dimensional
organization. Most microprocessor chips use a small number of layers of planar wiring. In contrast, neurons wire in a complete 3D fashion. Various studies have examined optimal geometric patterns of connectivity; and find that neurons, arranged in cortical columns, strike the optimum balance between two opposing tendencies: transmission speed and component density [39] [40]. Understanding the design rules underlying the micro-level column organization thus will be crucial for a better understanding of the brain as a whole.

Energy usage poses another conspicuous difference between neural networks in-silico and in-vivo. The Blue Gene/IP supercomputer used to model a cat’s cerebral cortex, requires several times more energy than an actual cat brain and still runs about 100 times slower [41]. The main issue is that transistor based networks convert most of the energy used into heat. Brains make considerably more efficient use of the energy they consume (about 20% of the resting metabolism) [42] [43] [44]. Of this energy, about 50% is used to drive signals along axons and across synapses with the remainder supporting the maintenance of resting membrane potentials and the vegetative function of neurons and glia. It should be noted that cortical gray matter consumes a higher proportion of energy (75%) than white matter, highlighted by the fact that global connectivity in the cortex is very sparse: the probability of any two neurons having a direct connection is around 1:100 for neurons in a vertical columns 1 mm in diameter but only 1:1,000,000 for distant neurons [45]. These findings point to another major design rule: use the majority of resources for modular, local signaling and computing (gray matter) and spend the rest on long-range communication (myelinated axons: white matter) linking the distributed computational units.

The development of a FCU promises several important applications for brain disorders as well. Crespi et al. (2009) demonstrate a diametric relationship between the genes affecting autism and schizophrenia. Specifically, there exist four specific genetic loci at which deletions are linked to autism and duplications to schizophrenia, or conversely, deletions are linked to schizophrenia and duplications to autism [46]. This unexpected structural link between two otherwise unrelated brain disorders has several important applications. First, it is indicative of a fundamental similarity between inherited brain disorders as a whole - not only in the way that they affect the brain, but the patterns of genetic affect that cause them. Second, it provides an important application for the FCU not only as a way to quantify what we define as cognition, but as a benchmark for faulty brain processes. Like cognition, language and DNA itself, brain disorders may be analogous in the sense that they are also complex combinations of a few simple processes such as genetic mutation at specific loci. Finally, energy constraints also govern the optimal coding schemes. It is well established that redundancy reduction is a major design principle in the cortex [47]. By maximizing the ratio between information coded and energy expended for signal transmission and resting state maintenance, sparse coding regimes improve energy efficiency [48] [49] [50]. Sparse coding regimes, in which a small proportion of neurons signal at any one time, have the additional benefit of a large representational capacity [51] [52].

An improved appreciation of the geometrical and energy constraints that are fundamental to the functioning of the brain should provide a valuable guide for the study of neuronal network architecture and coding. We have already learned that neurons do not statically integrate information as transistors do and also that the electric fields generated by neuronal activity in turn can affect neuronal activity itself [53]. We should thus ask whether it is appropriate to assume a binary code for neural networks or whether a more holistic view is required. Current attempts to computationally mimic brain function attempt to do so by fitting form to expected function; our unitary system proposal offers an alternative to persisting in a computer hardware-inspired understanding of the human brain by asserting the necessity of comprehending the physical expression cognition on each medium at which it operates.
8 Model

We are proposing a system that models the different mediums of brain function in a mathematically uniform manner. This system manifests itself in many ways relating to brain physiology, from neuronal activity, to molecular chirality and frequency oscillations. An improved appreciation of the geometrical and energy constraints that are fundamental to the functioning of the brain should provide a valuable guide for the study of neuronal network architecture and coding. A better understanding of intelligence can help scholars of artificial intelligence overcome the energy consumption and efficiency obstacles they face in reproducing the human brain’s performance at the hardware level. Modeling brain function promises insights to the paradox of the low energy consumption and high cognitive output of the brain. A method for efficient coding at the level of neuro-linguistics is explained here. The process by which the brain analyses the mental state of an individual provides an efficient learning algorithm and a great example of efficient coding. One of the ways in which the brain examines the mental state of an individual is to analyze the mind’s abstract structures. Throughout the brain there are various forms of activations (electrical, chemical, biological) each contributes individually or within groups to the formation of new concepts, which define a positive or negative mental state calculated by the Mood State Indicators algorithm (MSI) [34].

Brain areas that underlie value-based choices should receive value-based inputs from all choice options and then decide a winning value choice. Several authors used a neural network model to mimic the choice decisions and found that value-based decision-making occurs at low frequencies of <10Hz. They studied Magnetoencephalography (MEG) data from humans to confirm their network model and determine brain areas corresponding to the value comparison. The parietal and the prefrontal cortices were implicated in value-based choices. The advantage of MEG is that possesses high temporal resolution as the value comparison is being made and can detect more accurately than fMRI the time after the decision has been formed and how the choices are made over time [54].

The parameters that are used for the network model such as frequency oscillations are related to biophysical properties such as inhibitory (-) inputs, concentration of neurotransmitter, which can be measured locally using fMRI/MEG/magnetic resonance spectroscopy (MRS). Therefore MEG can be used to measure FCU in frequency oscillations associated and fMRI used to measure brain areas involved during axiology (another level of FCU analysis) using the MSI algorithm during patient testing in and MRS can be used to determine FCU in neurotransmitters involved. For example, the frequency oscillations that are based on excitation/inhibition balance in patient’s brain is measured using fMRI+MRS+MEG as patients are in discourse, and simultaneously, excitatory/inhibitory neurotransmitters are correlated. Therefore, multiple modalities can be used to measure FCU at different levels in parietal and prefrontal cortices in response to discourse, where mood states and word value can be also calculated.

By recording from the OFC and ACC of monkeys’ brains some scientists were able to train monkeys to make decisions. They found that OFC neurons evaluate a choice based on recent choice values, while ACC neurons integrate multiple decision parameters, positive and negative values, predictions and errors to evaluate choice. If these brain areas are damaged, it results in impaired decision making, suggesting these areas as the seat of value comparisons in the brain. The neuronal coding scheme employed by prefrontal cortical neurons was based on associating different firing rates to different aspects of decision-making. However, they found that ACC uses a positive valence-encoding scheme for all aspects of decision-making and evaluates reward prediction errors [55].

The OFC is likely to be activated during context based word value decisions and encodes an adaptive
value signal that changes based on recent value history, while ACC during and encoded value on a fixed scale to evaluate how choice outcomes are related to prior expectancies. This provides the subset of prefrontal cortical neurons that are involved in determining value and can be identified in imaging studies as the regions correlated to word value axiology. The coding scheme of neurons to calculate value and compare is reminiscent of FCU and comparing + and – values and forming an aggregate output. Hence, we can suggest that a write modality to feedback different word values or mood states can be used to control OFC and ACC to improve mood states and decision making.

Hence, if we were to trace the activations that trigger such behavioral outputs we find them at various levels. A series of such activations form an activation set, this set represents a connected structure for each activated region, which is defined in terms of a node. The node circumference changes based on the duration of the region being activated and the reflexivity is due to the re-activation of this region at different instances. Each node representing a form of activation can also be connected to another node that can vary in shape and time orientation. The segment connecting the various nodes to one another represents the time orientation. Once this connected structure is formed, a new activation set is created. This activation set can also be connected with other activation sets to create a concept set. In other words, nodes within the same set are added together based on the unary system computation and they are represented in terms of waveform signals that are weighted by a statistical coefficient to produce a resultant active node.

Figure 9 introduces the concept set. Concept sets consist of two or more activation sets. These sets have been generated through the interaction of neural activity at the various cognitive levels. Each node of those sets represents an active region that is either dependent on other regions to complete its “cognizant” state or is self-dependent but cannot form a comprehensive state. Two active states can form a uniform concept and interact together to define a concept set.

Each node represents an active region of the brain, which consumes energy to maintain continuous neural activity. Throughout the various stages of our analysis energy is being generated and propagated from the molecular level to every signal mechanism. Looking closely at the activation sets we can clearly distinguish the various levels of interaction ranging from chemical/biological, electrical to cognitive-energy activation of brain regions. Components at each level interact with one another in a unary structure producing a consistent function of activity. As a whole the various levels that activate these brain regions form a unitary system of operation that affects the brain from the molecular through biological to the cognitive levels to interaction.
The concept set is the set of concepts derived from the connected-activated sets, which produce an axiological value that represents a mood state after being projected on a positive and negative plane. The concept set is defined by the activation set with higher value. Therefore, a concept can be formed when other activation sets are elements of it, but their contribution (effect) is diminished by another dominant activation set. The value of the resultant concept set can have various ranges from positive infinity to negative infinity, only after its projection that these values become a unitary positive or negative. Correlated and dependent concept sets form a conceptual metaphor that remains consistent in terms of its axiological value despite changes in time.

While decoding the exact composition of each “letter” in the alphabet of thought from the sheer combinatoric complexity of the brain is outside the scope of this paper, we have shown biophysical evidence for a set of simple signals that compose cognition, and using the unitary system, we have also demonstrated a clear link between those signals and the linguistic expressions of conscious thought. In particular, we argue that concepts articulated linguistically can be placed on a continuum of positive and negative values based on factors such as environmental variables and linguistic and behavioral context.

We argue that these processes are closely interrelated and can be described mathematically in a uniform manner. The multi-level model of information exchange in biological systems, from abstract concepts in the brain down to the molecular interactions guides the development, functioning, and communication within the brain. The schema of coding is utilized to express behavior, intelligence, cognition, and conscience. The schema code governs fundamental neuronal communication across all brain activities, and forms the fundamental unit of thought. The link between cognition and linguistic articulation means that the unitary system can be used to model cognition both at the biophysical level and at the linguistic and system levels, and showed that it is imperative to comprehend its
expression on each of these mediums because of their mathematical continuity.

Combining the unitary system with empirical data promises to help determine not only the quantitative minimum of cognition (i.e., the number of active neuronal connections that form conscious thought), but its qualitative foundations as well; the DNA of brain language and human thoughts. The decomposition of cognition into these basic units governed by an intuitive mathematical model has potential applications in a number of fields, such as neurological and psychiatric therapy, discourse analysis, and a better understanding of intelligence that can help scholars of Artificial Intelligence overcome the energy consumption and efficiency obstacles they face in reproducing the human brain's performance at the hardware level.

The concept of mind-brain duality argues that the mind is fundamentally different from the brain in the way it functions and operates. We theorize that a Fundamental Code Unit (FCU) permeates the underlying workings of both the mind and brain's linguistics and communications systems, respectively. In this way, both the mind and the brain maintain schematics and axiomatic coding substrates similar to those found in DNA. Reconciling the paradox presented by the mind/brain duality concept promises to provide insight into several important phenomena, such as the location and composition of memories forth, as well as how they fit into the structure of the conscious mind. This work introduced a new “understanding” of cognition based on physical processes within the brain and its internal configuration at various modalities. Ongoing to this work we will be reporting experiment results from testing the Fundamental Code Unit (FCU) current results, as well set the foundation for future hypothesis found in contemporary research and the theories underlying it.

Acknowledgments

Special thanks to Ian Andrews, Deepti Rao, Sergey Kanareykin, Mark Atallah, Richard Wirt, Rebecca Fahlstrom, Michael Kohl and Adam Howard for Original Research support and edits. This work was enabled in part by The University of Oxford, U.S. Department of Defense and the Brain Sciences Foundation.

Author Biography

Newton Howard (nhmit@mit.edu) is the director of the Mind Machine Project and a resident scientist at the Massachusetts Institute of Technology (MIT). He received his Doctoral degree in Cognitive Informatics and Mathematics from La Sorbonne, France where he was also awarded the Habilitation a Diriger des Recherches for his leading work on the Physics of Cognition (PoC) and its applications to complex medical, economical and security equilibriums. While a graduate member of the Faculty of Mathematical Sciences at the University of Oxford, England, he proposed the Theory of Intention Awareness (IA), which made a significant impact on the design of command and control systems and information exchange systems at tactical operational and strategic levels. He has served as the Chairman of the Center for Advanced Defense Studies (CADS), the leading Washington, D.C, National Security Group and is currently its board director. He is a national security advisor to several U.S. Government organizations. He recently published The Mood State Indicators (MSI) algorithm, which modeled and explained the mental processes involved in human speech and writing, to predict emotional states. His natural-language approaches to systems understanding and design pointed to the multi-dimensional structures embedded and nested within speech-based cognitive systems, which have led to much advancement in building more accurate cognitive engines for modeling behavioral and cognitive feedback. Dr. Howard works with multi-disciplinary teams of
physicists, chemists, biologists, brain scientists, computer scientists, and engineers to reach a deeper understanding of the brain. Dr. Howard's research efforts aim to improve the quality of life for so many who suffer from degenerating conditions currently considered incurable. Advancing the field of brain sciences opens new opportunities for solving brain disorders and finding new means for developing artificial intelligence. Dr. Howard's most recent work focuses on the development of functional brain and neuron interfacing abilities. To better understand the structure and character of this information transfer he concentrated on theoretical mathematical models to represent the exchange of information inside the human brain. This work has proven applicable in the diagnosis and study of brain disorders and has aided in developing and implementing necessary pharmacological and therapeutic tools for physicians. He has also developed individualized strategies to incorporate solutions for psychiatric and brain prosthetics. Through collaborative research efforts with MIT and Boston University, Howard has studied principles of behavioral and molecular mechanisms affected by brain diseases.

References


Abstract
The significant disparity in computational efficiency between the biological and artificial brain (i.e., attempts to mimic brain function using contemporary computer hardware) suggests the need for a new model of human cognition. Based both on observed behavioral and linguistic patterns and on biochemical evidence, we propose a mathematical model for the structure of cognition based on unary mathematics. In particular, we argue that the temporal state transitions that govern conscious thought can be rigorously described using a set of functions we call the Unitary System. In addition, we populate this cognitive superstructure with units based on the Fundamental Code Unit (FCU), which specifies an indivisible unit of conscious thought over time at the quantum, biochemical and abstract concept levels. Ultimately, the Unitary System/FCU comprises an exercise in biomimetics. That is, we argue that artificial intelligence will benefit most by gathering insight from the structure and function of human intelligence, and a detailed understanding of neuronal function and network organization is required prior to neuronal network modeling attempts.

Keywords: axiology, activation set, energy consumption, unary system

1 Introduction
Over many generations, physiological developments tend to be guided by economy and efficiency [1]. Thus, achieving a more nuanced understanding of the design rules, both at the structural and unit levels, that govern brain function, is an important prerequisite to creating effective and useful models of the brain. As Ramon Y Cajal stated about a century ago, “all of the formations of the neuron and its various components are simply morphological adaptations governed by the laws of conservation for time, space and material”[2]. Neurons utilize significant quantities of energy to process information. Neuronal processing has been shown to comply with optimization models permitting the maximization of information intake. Energy is utilized in every active neural event and in order to maximize energy efficiently the structure of neurons in the brain is dense. A number of physical and structural differences between the brain and man-made computational hardware account for the vast efficiency gap. For instance, the brain has evolved a highly efficient neural code for the transfer of information over its highly specialized networks. In addition, it is organized in three dimensions (instead of two), and packs its networks into physically dense spaces.

Energy usage goes towards maintenance of the brain as well as anatomical productions. Contemporary hardware such as microprocessors generally use a few layers of planar wiring, but that is the extent of its three-dimensionality. Neurons, however, are
arranged in columns and are inherently more efficient as a result. They are both faster in information transmission and denser [3-5]. Consequently, a better understanding of this structural makeup has significant implications for all fields that use information processing.

Another expensive energy cost is due to long-range signal processing in the brain. As axons take up a large space in the brain, axonal transmission across long ranges utilize large amounts of energy as they dissipate energy during transmission. However, global connectivity is less and thus reduced volume of occupied by long-range connections.

The topographic organization of sensory cortical areas into columns provides an optimal geometry for energy efficient processing. The neurons in the columns are connected to each other at a much higher rate than neurons farther apart. Thus cortical columns maximize axon length. In addition cortical columns also have dense dendritic and axon branching thus enhancing area for input and output and minimizes transmission delays. However, this local connectivity and processing is not well communicated at distant sites, and only a small number of neurons that have long range connections are able to communicate this information. To overcome this neurons use synchronous firing between the connected areas to be able to communicate local information globally.

The Blue Gene/P supercomputer, built by IBM, was recently used to mimic a cat's cerebral cortex by simulating the number of connections in the latter. However, it requires significantly more energy and is slower by a factor of approximately 100. [18]. This is due to a number of factors, such as the fact that transistor-based networks convert most of the energy they use into waste heat. By contrast, the brain makes more efficiency use of energy – about 50% is used to power signaling via axons and synapses, and the remainder supports the resting membrane potentials and vegetative neuron and glia function [6-8]. About 1% of neurons in 1mm diameter vertical columns are directly connected, but only .000001% of distant neuron pairs are directly connected. This suggests that local signaling comprises the bulk of energy use, and the linking distributed computational units account for the rest. In particular, this configuration leads to optimal energy efficiency [9].

A large percentage of energy usage is taken up by synaptic transmission between neurons, primarily associated with the Na+/K- pumps but minimally utilized by neurotransmitter and signaling molecule recycling. Energy is required also to maintain neurons at their resting membrane potential. Laughlin (2001) showed that action potential propagation and postsynaptic current utilize the largest amount of energy in the cortex.

The amount of energy used is correlated to neural performance. Neural performance can be measured by signal to noise ratio and speed of response. To optimize the energy efficient neural coding process it is important to increase the information coded while decreasing the energy involved. In terms of spiking activity, the goal is to enhance the amount of information gained from a train of action potentials during a limited period of time. Another energy optimization technique is sparse coding where maximal information is gained from a small number of active neurons as opposed to a large population. The optimal neural code enhances the information gained from a small number of active neurons, i.e. minimal resource use for enhanced performance. Another energy efficiency issue is geometric constraints on neuronal packing and wiring. Reducing redundancy is a major guiding principle in the design of the cortex [10]. By reducing energy expended and simultaneously increasing the amount of information coded, so-called sparse coding regimes increase the energy efficiency of the brain and allow for a larger representational capacity [11-15].

The bandwidth of neurons is higher than would be expected from sparse coding. One of the main energy constraints in the brain is noise, for example caused by ion channels or signaling molecules. The noise and unreliability of neural signals is counteracted by plasticity mechanisms that make synaptic transmission more reliable and energy efficient. We know already that neurons do not integrate information in the same manner as transistors, and
that the electromagnetic fields that are a by-product of neurological activity can in turn affect the outcome of that activity [16]. Consequently, the question arises as to whether a binary code, like that used in transistor information processing, is appropriate for neural networks.

The system we propose, called the unitary system, manifests itself in many ways relating to brain physiology, from neuronal activity, to molecular chirality and frequency oscillations. To best understand the unitary system it is imperative to comprehend its expression on each of these mediums. Various atoms each have a unique atomic marker that defines its behavior and interactions with other atoms. Interactions among neuronal molecule cause the changes in ionization energy that ultimately leads to synaptic exchange. These differences are the result of particle configuration and, most importantly, the ionization energy that is determined by the atom’s valence [17]. Thus, when atoms approach one another, it is their outer shells that are responsible for chemical reactions such as synaptic exchange.

Changing chemical gradients within the brain, specifically those created by ions, create an energy and charge flow between neurons. This flow, characterized as action potentials throughout the brain, give rise to an energetic field that we call consciousness.

Some researchers have attempted to reverse-engineer the brain as a kind of computing system. IBM and Stanford University researchers have modeled a cat’s cerebral cortex using the Blue Gene/ IP supercomputer, which currently ranks as the fourth most powerful supercomputer in the world [18-20]. This machine uses 144 terabytes of RAM. Nevertheless, the simulated cat brain still runs about 100 times slower than the real thing [18]. In fact, using just 30 watts of electricity (i.e., enough to power a dim light bulb), our brains outperform the Blue Gene computer by a factor of a million. Consequently, several computer scientists maintain that a robot processor as smart as the human brain would require at least 10 megawatts to operate (i.e. the amount of energy produced by a small hydroelectric plant).

Despite all these efforts the brain remains a “three pound enigma” and our understanding of the phenomena listed remains rudimentary. Contrary to popular belief, silicon-based simulation is clearly not an adequate approach to deciphering the mind and brain; the energy efficiency metrics point at radically different levels at which brain’s computation takes place. Therefore, we need to look at brain-specific information exchange models, from cognitive to molecular, to understand the mechanisms that underlie cognition and awareness.

2 Model

In order to solve the problems facing scholars of Artificial Intelligence that we have laid out, we propose a system that models the different mediums of brain function in a way that can be applied in-silico as well as in-vivo. This system is not limited to a single level of analysis, and encompasses molecular chirality, frequency oscillations, as well as fluctuations at the quantum level. A better understanding of the geometrical and energy factors that are fundamental to brain function promises valuable insight for studying neuronal network architecture. Our model promises efficient coding at the neurolinguistic level. As an example, we turn to the process by which the brain ascertains the mental state of an individual, since it is an example of an efficient learning algorithm. One way that this is done is to analyze the mind’s abstract structures. Throughout the brain there are various forms of activation (electrical, chemical, biological). Each contributes individually or in groups to the formation of new concepts, which define a positive or negative mental state calculated by the MSI algorithm [19].

Series of activations form an activation set. This set represents a connected structure for each activated region, which is represented as a node. The node circumference changes based on the time frame of the active region and the reflexivity is due to the re-activation of this region at different instances. Each node representing activation forms can also be connected to another node that varies in orientation
and time frame. The segment connecting nodes to one another represents the time orientation. Once this structure is formed, a new activation set is created. This activation set can also be connected with others to create a concept set. In other words, nodes within the same set are added together based on the Unitary System computation and they are represented as of waveform signals. These in turn are weighted by a statistical coefficient to produce a resultant active node [20].

Figure 1 above shows two different concept sets {1, 2}. Each of these consists of two or more activation sets. Activation sets such as these are generated through the interaction of neural activity at various cognitive levels. Each node in those sets represents an active region that is itself in one of two states. First, it may be dependent on other active regions consisting of other sets to complete its “cognizant” state. Alternatively, it can be self-dependent but cannot form a comprehensive notion. Two active states that are self-dependent can form a uniform concept couple to define a concept set.

The concept set is derived from connected-activated sets. The value of each of these represents a positive or negative mood state. The resulting concept set after formation by being diminished by another activation set can have any positive or negative value. Only after it is projected do these values assume a unitary character.

For example, if we were to think of “Sky” as a concept set and in order for us to determine what it’s axiological value is we need to follow the current procedure. The concept “Sky” is a resultant of various activation sets in which each contribution was based on the level of its activation within a specific time frame and orientation. If we were to consider another concept such as “Dark” most likely the same computation applies but a negative axiological value would be assigned. At another point in time and according to the same patient the concept “Morning” can be formed based on these two concepts and it will result with a positive
Axiology since a “Dark Sky” might resemble the end of a day and hence announcing the coming of the “Morning”. This axiological value is valid within these time-division constraints and it might change if the constraints were to change, similarly for other patients. The use of a learning algorithm will thus enhance our computational accuracy while reducing evaluated errors for specific individuals.

Hence, we will introduce a mathematical framework for Unary Topological Axiologies describing mental states. Its mathematical structure can be defined in terms of instantaneous and continuous representations. For instance, this computation from can be applied:

1. Begin with a set $S$ (uncountably infinite) representing brain regions, which may be activated by some means. We introduce a $\sigma$-algebra $A$ on this set, and call the elements $a \in A$ activation sets (by definition $a \subset S$). Now introduce a second set $W$ whose elements are labeled concepts in the brain, which correspond to words. For some subset $\mathcal{A} \subset A$ there is a mapping $P: a' \in \mathcal{A} \rightarrow \omega \in W$ called the concept activation mapping. The elements $a'$ of $\mathcal{A}$ are action potentials. Let $P^*: \omega \in W \rightarrow \tilde{a} \in \tilde{A}$ be a mapping we call the brain activation mapping. Let $\mu$ be a measure of $S$, and let $F: A \rightarrow \{+, -\}$ be a parity mapping. All axiology is mapping $\Xi: W \rightarrow \{+, -\}$ generated by computing

$$f(\omega) = \int_a F(s) d\mu \quad (1)$$

with

$$a = \tilde{P}(\omega) \quad (2)$$

and then projecting

$$\Xi(\omega) = \text{sign}(f) \quad (3)$$

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>brain regions</td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>activation sets</td>
<td>$a \in A \Rightarrow a \subset S$</td>
</tr>
<tr>
<td>$\mathcal{A}$</td>
<td>concept activation sets</td>
<td>$\mathcal{A} \subset A$</td>
</tr>
<tr>
<td>$W$</td>
<td>concepts</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>concept activation mapping</td>
<td>$P: a' \in \mathcal{A} \rightarrow \omega \in W$</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>axiology</td>
<td>$\Xi: W \rightarrow {+, -}$</td>
</tr>
<tr>
<td>$F$</td>
<td>parity mapping</td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>weight mapping</td>
<td></td>
</tr>
</tbody>
</table>

Various experimental data on the distribution of energy-consumption among known cellular mechanisms and their contribution to total usage have been performed. Many conclusions showed that for mammalian brain, 50% of the total energy consumption is associated with signaling, and recent studies in NMR spectroscopy, which associate energy usage with neural function by following the turnover of identified metabolites and neurotransmitters, suggest that signaling
accounts for 80% of the total consumption in cortex [21]. This signaling cost is dominated by the activity of the Na+ /K+ pumps, which, like mitochondria, are concentrated in axons and at synapses. The signaling cost of recycling second messengers and neurotransmitters is less than 10%.

Taking biophysical and biochemical measurements of membrane conductance and transmitter release, one calculates the numbers of ions and molecules that are used to generate signals in known components such as synapses, dendrites and axons. These numbers define energy usage in terms of the ATP that must be hydrolyzed to maintain ionic concentration gradients and recycle signaling molecules [20]. Basic biophysics links energy usage to two fundamental measures of signal quality: signal-to-noise ratio (SNR) and bandwidth (a measure of speed of response). Extra energy is required to improve the SNR because reliability increases as the square root of the number of stochastic events that are used to generate signals. Each stochastic signaling event, such as opening an ion channel or releasing a synaptic vesicle, requires extra energy. Raising the bandwidth also requires extra energy. The membrane time constant must be reduced by increasing conductance and this increases the flow of ionic current. Hence, minimizing the SNR and bandwidth will save energy [Ibid].

Energy-efficient neural codes increase the ratio between the information coded by neurons and the energy involved. The ability to represent information depends on the number of recognizably different patterns of spikes that can be generated in a useful time interval. This number — the ‘representational capacity’ — is a fundamental measure of neural performance [Ibid]. It defines the number of situations and actions that neurons can specify, and is closely related to the number of bits coded.

Sparse coding, in which information is represented by the activity of a small proportion of neurons in a population, greatly increases energy efficiency. The optimum distribution of activity depends on the ratio between the signaling cost of making a spike and the fixed cost of maintaining a neuron in readiness to spike. When spikes are expensive and the fixed cost is negligible, a ‘grandmother neuron’ code is cheapest [Ibid]. This code associates each unique event with a single active neuron.

One method of energy efficient coding at the cellular level in the brain is explained here. The method by which the brain processes information through the integrating information from sources, firing action potentials, provides an efficient coding system. The brain is bombarded on a daily basis with a barrage of incoming information from different sensory areas and inputs. The most energy efficient method to code information is for neurons to be activated only when they receive important information. One way to achieve this is to transmit information by firing action potentials only for important inputs and not for every input that comes into the brain. In this way, irrelevant information is not coded. Neurons receive inputs from different sources and each of the inputs carry different weights, translated by the size of neuronal depolarization. If inputs are strong enough in size, i.e. for important information, neurons are depolarized sufficiently to reach firing threshold and thus fire an action potential that is transmitted in the network. If the input is not strong enough, i.e. for unimportant information, neurons do not reach threshold and do not fire action potentials. Another efficient coding system in an individual action potential is neuronal repolarization. Once neurons fire an action potential, the membrane potential reverts to its original value due to the ion outflow [20]. This allows for new incoming information to be coded in the brain as neurons are ready to be depolarized again and fire the next set of action potentials. One of the efficient coding systems exists at the neuromuscular junction. At the neuromuscular junction, virtually every action potential in the presynaptic motor neuron triggers an end plate potential in the postsynaptic muscle cell. In this manner, none of the energy utilized to fire an action potential is wasted, unlike at central synapses. At central synapses, the postsynaptic neuron commonly receives signal from many...
presynaptic neurons. A single neuron can be affected simultaneously by signals received at multiple excitatory and inhibitory synapses.

Another method of energy efficient coding at the network level is explained here. As Cajal suggested, the structure of a neuron is related to its functional role. Axons, dendrites and dendritic spines are all important structures for increasing processing capabilities of a neuron. Cortical columns in the brain are optimized to process specific input features into the brain and are designed to enhance efficient coding by the way neurons are wired in the column [21]. The goal of energy efficient brain processing is for output information being equal to the input that goes into the brain and to connect the two streams of information, output and input while utilizing minimal anatomical space.

Another method for energy efficient coding at level of neural oscillations is explained here. One can deduce neuronal brain states and brain regions involved in different mood states by measuring neural oscillations. Neural oscillations determine brain function and different oscillations have different functions in different brain regions.

Oscillations range from low delta, theta, alpha, beta and high gamma. Delta waves <4Hz are important for low level arousal activity. Theta waves are 4-8Hz waves are important for creative and intuitive thinking. Alpha waves 8-12Hz are important for meditation and relaxation. Beta waves 13-30Hz are important for mental activity processing. Gamma waves >30Hz are important for high-level information processing. Changes in normal neural oscillations can reflect neurological disorders in the brain. Thus it is imperative to understand the nature of such frequencies, their causes, their ranges, and the relation of each range to each disorder. As specific oscillations signatures are associated with specific mood states, the energy utilized for a specific mood state can be measured. Energy utilized by neural oscillations can be measured in terms of energy output in single ATP transactions during a single wave in an oscillation and thus correlated to the corresponding mood state. An example of an energy expensive mood state would be a state that scores high – in the MSI algorithm such as depression while an example of a inexpensive mood state would be the state that scores 0. This suggests that in order to maintain low energy during high cognitive tasks, it is better to keep mood states open and aware in stressful scenarios as to not lose energy.

Chronic pain is a debilitating condition that could be due to neuropathic or nociceptive conditions. There does not exist direct objective neural correlates to pain in the brain. One study identified a neural signature of brain wave activity in 8-12Hz in the sensory thalamus and periaqueductal gray (PAG) and that is significantly correlated to subjective neuropathic pain intensity. Thus study further
demonstrated that using deep brain stimulation and local field potential analysis an increase in amplitude of 10Hz neural oscillations was observed in both sensory thalamus and PAG. Green et al. (1974) showed that in the power spectra there was an increase in 8-12Hz activity in the PAG and an increase in the 17-30Hz activity in sensory thalamus. Treatment for chronic pain can be brought about by neurofeedback. Once the abnormal increase in 8-12Hz and 17-30Hz is detected in the PAG and sensory thalamus, neurofeedback via transcranial magnetic stimulation (TMS) or ultrasound can be employed to decrease and eliminate the abnormal oscillations. It is also important to determine the cause of this abnormal enhancement in 8-12Hz activities and why this increase results in perception of chronic pain. The causes of altered oscillation might be due to changes in neurotransmitter concentration or availability, ion channel changes.

4 Conclusion

The human brain differs fundamentally from current silicon-based attempts to mimic it on a number of levels. Apart from the fact that neurons are organized and networked differently from transistors and other man-made computational hardware, there exist several inherent differences at the unit level between in-vivo and in-silico computational examples. In this paper, we have argued that these fundamental differences account for the staggering superiority of the brain to date, both in terms of cognitive performance and energy efficiency. In addition, the optimal coding schemes for the neural networks in the brain that drive cognition make highly dynamic use of energy and cellular resources based on the qualities of the information being conveyed, thus enabling optimization on both the hardware and configuration levels. Specifically, sparse coding regimes that capitalize on a small number of active neurons instead of activating entire networks, as well as the redundancy-minimizing connectivity and placement of neurons, both contribute to the computational advantage of the human brain.

We take this conclusion a step further by proposing that in order to successfully close this gap, the mathematical characteristics of neurological networks and coding schemes must be first understood and then made the basis for future hardware-based computational cognition. In our case study of chronic pain, we apply the specific brain wave oscillation patterns identified by Green's 1974 study, to the specific mechanisms that drive this perception – neurotransmitter concentration and, potentially, ion channel alterations. This phenomenon provides a clear and intuitive stimulus/response feedback system against which to test various neurological mechanisms for their effects on the brain, and is a prime example of the type of inference that is needed in order to determine its deeper cognitive characteristics.

Acknowledgments

Special thanks to Ian Andrews, Deepti Rao, Sergey Kanareykin, Mark Atallah, Richard Wirt, Rebecca Fahlstrom, Michael Kohl and Adam Howard. for original research support and edits. This work was enabled in part by The University of Oxford, U.S. Department of Defense and the Brain Sciences Foundation.

Author Biography

Newton Howard (nhmit@mit.edu) is the director of the Mind Machine Project and a resident scientist at the Massachusetts Institute of Technology (MIT). He received his Doctoral degree in Cognitive Informatics and Mathematics from La Sorbonne, France where he was also awarded the Habilitation a Diriger des Recherches for his leading work on the Physics of Cognition (PoC) and its applications to complex medical, economical and security equilibriums. While a graduate member of the Faculty of Mathematical Sciences at the University of Oxford, England, he proposed the Theory of Intention Awareness (IA), which made a significant impact on the design of command and control systems and information exchange systems at
tactical operational and strategic levels. He has served as the Chairman of the Center for Advanced Defense Studies (CADS), the leading Washington, D.C. National Security Group and is currently its board director. He is a national security advisor to several U.S. Government organizations. Howard recently published The Mood State Indicators (MSI) algorithm, which modeled and explained the mental processes involved in human speech and writing, to predict emotional states. His natural-language approaches to systems understanding and design pointed to the multi-dimensional structures embedded and nested within speech-based cognitive systems, which have led to much advancement in building more accurate cognitive engines for modeling behavioral and cognitive feedback. Dr. Howard works with multi-disciplinary teams of physicists, chemists, biologists, brain scientists, computer scientists, and engineers to reach a deeper understanding of the brain. Dr. Howard's research efforts aim to improve the quality of life for so many who suffer from degenerating conditions currently considered incurable. Advancing the field of brain sciences opens new opportunities for solving brain disorders and finding new means for developing artificial intelligence. Dr. Howard's most recent work focuses on the development of functional brain and neuron interfacing abilities. To better understand the structure and character of this information transfer he concentrated on theoretical mathematical models to represent the exchange of information inside the human brain. This work has proven applicable in the diagnosis and study of brain disorders and has aided in developing and implementing necessary pharmacological and therapeutic tools for physicians. He has also developed individualized strategies to incorporate solutions for psychiatric and brain prosthetics. Through collaborative research efforts with MIT and Boston University, Howard has studied principles of behavioral and molecular mechanisms affected by brain disease

References


Cognitive Architecture: INTEGRATING SITUATION AWARENESS AND INTENTION AWARENESS

Newton Howard
Massachusetts Institute of Technology, Cambridge, MA 02139
newton.howard@brainsciences.org nhmit@mit.edu

Abstract
Systems that incorporate situational awareness or enhance that of human operators have found successful applications, especially those that deal with high volumes of data. However, such applications give the impression that situation awareness is complete when only a few parameters are with an emphasis on raw quantitative data, such as position and speed of objects. These parameters, however, tell us little about situations when they are governed by human actors, and hence by the complex human mind. A better understanding of intention enables us to overcome that obstacle. By integrating concepts of intention awareness, or internal characteristics of actors themselves, it is possible to significantly improve situationally aware systems.

Keywords: intention awareness, situation awareness.

1 Introduction

Parking a vehicle in a busy parking lot requires awareness and safe navigation. Such a daily routine task requires the driver to quickly grow aware of the location of parking signs, other cars, and people about the lot. With a brief glance around, the driver uses these objects and their relative positions to represent his current situation. The physical distribution of objects in time and space dictates how a driver accomplishes the task of parking a vehicle – the Parking Dilemma [11].

In order to successfully navigate through a field of static and moving objects, the driver must develop awareness not only of the situation, but also of other drivers’ and pedestrians’ intentions. Acts and motions will otherwise appear random. Some cars will suddenly change their course to turn into an open parking spot; others will reverse unexpectedly to vacate a spot. Drivers understand these motivations and compensate for them. Specific functions of automobiles such as turn signals assist in conveying intentions, effectively transferring manifesting an internal desire into the external physical realm, and augmenting awareness of other variables such as the car’s position, speed and direction.

Human activities are constrained by time, space and resources. A routine human activity, such as parking one’s car, requires integrating situation awareness and intention awareness if “time and resources” are limited. Assuming a changing environment, one can alternatively continue navigating until an empty parking space is found. Good drivers are sensitive to potential situations as they navigate through the lot, though situational awareness alone does not suffice to explain what makes one driver better or worse than another. In this example, it is clear that though situation
The Brain Sciences Journal

awareness alone would provide a clear snapshot of the current situation, a clear understanding of the intentions of other drivers is equally important to ensure a safe and successful navigation. Situational awareness only provides data for the present moment; operating without any intention awareness is akin to a nearsighted person driving down the expressway without glasses. Being able to see only 10-20 feet ahead at any given moment provides immediate situational awareness but does not sufficiently inform the driver of the road ahead. The faster he drives the more significant this distant vision and awareness become.

In a game of chess, situation awareness allows a player to make legal moves but will not necessarily enable him to understand the opponent’s strategy: You can play and you will probably lose. Intention awareness, on the other hand, will guarantee that the opponent’s strategy is understood but it won’t guarantee winning. You may understand the game and still lose.

This paper splits the discussion of situational awareness into four additional sections. In the first of these, the author conducts a literature review of relevant theoretical work in the field of situational awareness in order to respond to their strengths, weaknesses and counterarguments. We apply these results both in the second section, which frames the discussion of situational awareness by providing a working definition, and the third section, which demonstrates the utility of intentional awareness study in complementing applications of situational awareness. The conclusion synopsizes these findings and sketches out areas for future work in the field.

2 Alternative Theories and Approaches

2.1 The Zeigarnik effect

Bluma Zeigarnik, a 20th Century Soviet psychologist, discovered that human memory has selective storage properties; that is, information retention varies with its relevance to tasks immediately at hand [38]. Waiters, for instance, might memorize several simultaneous orders without error and not be able to recall them as soon as the orders were fulfilled. Consequently, this became known as the Zeigarnik Effect, which carries with it several important implications. First, learning is an expansive cognitive activity, and the more tension surrounds a particular task, the more one is likely to make progress in learning it. In addition, there is an amount of tension that creates the ideal learning environment, but excess tension hinders the process.

Zeigarnik conducted critical experiments about the recall of finished and unfinished actions to measure human association and “will power” during the years 1924 and 1926. Zeigarnik concluded that the “effect of intention is equivalent to the creation of an inner personal tension” [1]. Tension in physics refers to the state of one system relative to the state of surrounding systems. It is merely a manifestation of the fundamental forces that exists and cause change (motion) in the direction of equalizing the state of neighboring systems. This conclusion assumes a model of the mind as a dynamic system with subsystems, forces and force-fields. It also “presupposes a geometric representation of the person and a distinction of functional sub parts or “systems’ within the person, with a definite position in regard to each other” [1].

Zeigarnik’s theory contains three basic assumptions:

- **A1:** The intention to reach a certain goal G (to carry out an action leading to G) corresponds to a tension (t) in a certain system S(G) within the person so that t(S(G)) > 0. This assumption coordinates a dynamic model (system in tension) with “intention.”
- **A2:** The tension system t(S(G)) is released if the goal G is reached. T(S(G)) = if P accomplishes G. Zeigarnik uses the tendency to recall activities as a symptom for the existence of tension. The expectation of the existence of such a system is based on the final assumption:
- **A3:** To a need for G corresponds a force f(P,G) action upon the person and causing a tendency of locomotion toward G. If t(S(G)) > 0 then f(P,G) > 0. This assumption determines the
relation between need and locomotion - motion from one place to another place. In other words it means a construct of tension in the person and the construct of force for locomotion in the environment.

• A3a: A need leads not only to a tendency of actual locomotion towards the goal region but also to thinking about this type of activity; in other words, the force $f(P,G)$ exits not only on the level of doing (reality) but also on the level of thinking (reality); if $t(S(G)) > 0$ $f(P,R) > 0$ where $R$ means recall.

From the three assumptions $A1$, $A2$, $A3$, and $A3a$ it follows: the tendency to recall interrupted activities should be greater than the tendency to recall finished ones. This derivation can be made as follows. We indicate the completed task by $C$, the unfinished one by $U$, and the corresponding systems by $S(C)$, and $S(U)$ respectively.

We can then state:

(1) $t(S(U)) > 0$ according to $A1$
(2) $t(S(C)) = 0$ according to $A2$

Hence,

(3) $f(P,U) > f(P,C)$

according to $A3a$, on the level of thinking. In other words: there is a greater tendency to recall spontaneously unfinished tasks than finished tasks.

Zeigarnik’s main result was, $RU/RC = 1.9$, where $RU$ is recalled unfinished tasks, $RC$ recalled completed tasks.

Experiments where certain tasks were first interrupted but later on allowed to be finished served to prove that it is not the experiences connected with the interruption itself which are the cause of this result but the reaching or not reaching of the goal. In this experiment the recall was not more frequent than in the case of tasks completed without interruption. Zeigarnik expressed the basic relation between neighboring tension systems in the following way: $C1$: If $t(S1) \neq t(S2)$ and $B(S1) \times B(S2) \neq 0$, a tendency exists to change so that $t(S1) = t(S2)$, where $B(S1)$ and $B(S2)$ indicate the boundaries of the systems $S1$ and $S2$, $B(S1) \times B(S2)$ is their common part. This model presupposes definite assumptions as to the dynamic character of this field, e.g., if the systems corresponding to different needs or quasi-needs should be able to maintain different amounts of tension during a certain period, one will have to assume that this field is not too fluid. If it should be a very fluid field, any differences between the tension levels of the various systems would be found to disappear in a very short time because of the fact that the tendency of equalization resulting from the local tensions would not meet any resistance; in other words, if a quasi-need is coordinated to a tension system which may show its effect even over a considerable time interval, one has to assume that dynamically the person cannot be considered as entirely fluid. On the other hand, a person cannot be regarded as entirely rigid. Otherwise, the effect which one need has on other needs and on the tension level of the person as a whole could not be accounted for. A person, therefore, has to be conceived of as having a medium degree of fluidity in regard to the intercommunication of his tension systems. It is clear that this degree of fluidity may vary from person to person and from situation to situation for a single person. Assuming the constancy of the structural relations of a given set of systems (and assuming a temporarily impermeable boundary surrounding the set as a whole), one can express this statement in the following way: $C2$: Let us indicate the absolute difference between the tension $t(S1)$ and the tension $t(S2)$ of two neighboring systems $S1$ and $S2$ at the time the tensions are being built up by $|t(S1) - t(S2)|$ the time since then elapsed by $T$, the tension difference at this time by $|t(S1) - t(S2)|_T$, and the fluidity by $f$. Then we can state $F(T, f) = |t(S1) - t(S2)|_T - |t(S1) - t(S2)|_T$, where symbolizes a monotonically increasing function. This means: the change in the tension difference of neighboring systems depends upon the time interval and the fluidity. Of course, this holds true only if the tensions of these systems are not changed by other factors, e.g. release of tension by reaching the goal.
C1 and C2 are necessary conceptual elements of the construct tension. The coordination of needs and quasi-needs to this construct tension, therefore, makes it possible to derive a number of facts which may seem rather remote from the problem primarily investigated. These predictions could hardly be made without this specific dynamic theory, and therefore if they can be proved, they are of particular value for the confirmation of the theory. The difference in tension between systems corresponding to unfinished and finished tasks decreases with the time interval elapsed since the creation of the tension system.

Field theory maintains that behavior does not depend on the past nor on the future, but does depend on the present field [1]. This present field has a certain time-depth. It includes the “psychological past,” “psychological present,” and “psychological future,” which constitute one of the dimensions of the life space existing at a given time. This contrasts the philosophy of teleology that the future is the cause of behavior, and the beliefs of associationism that the past is the cause of behavior. Furthermore, it is an error to consider the assumption of directed factors as characteristic of teleology. The causal explanations in physics certainly do not avoid such assumptions: physical force is a directed entity, a vector. Psychology, too, becomes in no way metaphysical by resorting to constructs of vectorial character such as psychological forces. This permits a direct attack on the problems of directed action. In addition, by defining direction in terms of hodological space, an adequate representation is possible of what has been meaningful in some of the other claims of teleology. The puzzling relation between knowledge and dynamics which had a mystical character in teleology is made understandable at least in one fundamental point: it becomes clear why lack of knowledge has the effect of a barrier. The mysterious ability of animals to make roundabout routes can be rationally related to the fact that equilibria in the hodological space depend upon the totality of relations in the field. Human behavior is manifested through activities. Activities are performed by agents via tasks. Each agent has two models of the world; perceived reality model, and ontological reality model. Both models are mental states. Perceived reality is a present model of the world (beliefs, facts, uncertainties, information, the possible and the impossible, etc.).
is a system of temporal rules, variables, and relations that regulate, govern the world. Ontological reality is anticipated or desired model of the world. Further, the agent has a situation awareness – the agent’s situation in relation to his/her perceived reality. To simplify discussion, we will refer to perceived reality model as model of type A, and ontological reality model as model of type B. At the same time, each agent has a system of intentions that in effect defines the ontological reality model.

In short, type A model is the present model of the world. Type B model is a model of the world to be pursued. There is a system of intentions that defines a plan from type-A-model to type-B-model, and finally, there is situation awareness of the type-A-model. There is no situation awareness of type-B model yet since type-B model is a state of mind that has not been reached (has not been “booted” on the “operating system” yet). However, there is an intention awareness of type-B model which is in effect a plan that links type-A model to type-B model.

The fact that each agent at time (t) has two models, situation awareness and intention awareness, creates many challenges to highlight, such as multiple perceptions due to agents’ type-A models, multiple expectations due to agent’s type-B models, multiple self-situation awareness and situation awareness, and multiple self-intention awareness and intention awareness.

Hypotheses:
1. Model A(X) – Model A'(X) \(\neq 0\). Agent X type-A model at time t1 cannot be equal to Agent X type-A model at Time 2.
2. Model B(X) – Model B'(X) \(\neq 0\). Agent X type-B model at time t1 cannot be equal to Agent X type-B model at Time 2.
3. Model A(X) – Model A(Y) \(\neq 0\). Agent X type-A model at time t1 cannot be equal to Agent Y type-A model at Time 2.
4. Model B(X) – Model B(Y) \(\neq 0\). Agent X type-B model at time t1 cannot be equal to Agent Y type-B model at Time 2.

2.2 Popper’s Three World Model

According to Popper, every thing that exists can be placed in one of three worlds [42]. World 1 is the world of physical objects and states (inorganic, biology, and artifacts). World 2 is the world of states of consciences and subjective knowledge and experiences (perception, thinking, emotion, memories, dreams, etc.) World 3 is the world of culture created by man including all objective knowledge (cultural heritage, language, theoretical systems, etc.) The representation of the three worlds compromise all existence and experiences according to Popper.

Popper’s work is related to the theories on how the mind and the brain interact. It is argued that
the world of self-conscious, which is World 2, of an individual is developed based on the relationship with World 3 and the influence that World 3 exerts on the self. Popper has the hypothesis that the self-conscious mind is an “independent entity.” The self-conscious mind exercises a controlling role upon the neural events through a two-way interaction across an interface between World 1 and World 2. Knowledge in World 3 (culture and heritage) can be documented on various objects in World 1 (machines, books, pictures, etc.), and by that, creating a two-way interaction between World 3 and World 1. Therefore, World 3 has a two-way interaction with World 1 which has a two-way interaction with World 2. The interaction between World 2 and World 3 happens with World 1 as the mediator.

Popper also introduces the information flow diagram between the mind and the brain and the interactions between them [42]. The outer sense is related to perceptions given by the inputs of the sense organs (visual, auditory, tactile, smell, taste, etc.). The inner sense is compromised of many cognitive experiences (thoughts, memories, intentions, feelings, emotions, etc.). In the center of World 2 lies the ego that is the basis of the personal identity. There is also an interface that lines of communication between World 1 and World 2 occur, i.e. from the liaison brain to and from the components in World 2.

A few points have to be discussed regarding Popper’s model of the interaction between the brain and the mind. First, which part has a bigger control over the other in World 2, is it the outer sense or the inner sense. The second comment regarding the model is that some vague words such as Ego, the Self, and Will are used without reaching a universal definition that will help the reader understand the model. Without having standard definitions, a conflict may occur in the understanding of the model. The third comment is that it is seen that there is only a direct interaction going from the direction of the outer sense to the inner sense. There is no information flow from the inner sense to the outer sense. The only way is to have an indirect path though the Ego, into the brain in World 1, and back to the outer sense. The question is should information that has to flow from the inner sense to the outer sense always have to pass through the ego and the brain before reaching the outer sense? Or can there be a direct flow of information? We would conclude that the relation between neurons and the mind are still hard to prove.

2.3 Newell’s Criteria

Newell proposed a list of criteria that the human cognitive architecture has to satisfy in order to be functional [43]. One may see that the criteria are broad and it might be hard to evaluate any theory based on these criteria. The criteria can be condensed into 12 criteria. Newell's functional criteria for a human cognitive architecture are:

1. Flexible behavior: Behave as an almost random function.
2. Real time: Operate in real time.
3. Adaptive behavior: Exhibit rational and effective adaptive behavior.
4. Vast knowledge base: Use vast amounts of knowledge about the environment to affect performance.
5. Dynamic behavior: Behave robustly during error, the unexpected, and unknown.
9. Learning: To learn from environment
11. Evolution.
12. Brain: Be realized within the brain.

We are also introduced to Connectionism. Connectionism is considered to be the computationally modern successor to behaviorism. Connectionism focuses on learning and emphasizes how behavior can arise from adaptive response to experience (Newell's criteria 3 and 9). It rejects any abstractions (Newell's criterion 6) unless it is a matter of verbal behavior (Newell's criterion 8). Even though behaviorism rejects consciousness, connectionism addresses it.

The ACT-R theory, which is a theory of cognition, has been put to the Newell test by applying it to a wide range of subjects. The ACT-R architecture is an abstraction from the neural level. The model contains components: goal, declarative memory, perceptual, and motor modules. The flow of cognition in the system is from the response to the current goal, active information from declarative memory, information in perceptual modules (both vision and audition), and the state of the motor modules.

Both connectionism and ACT-R model have low scores when it comes with criteria 8 (self-awareness and consciousness) and 11 (evolution). ACT-R focuses on symbolic structures and processes and offers explanation that links to empirical behavior while connectionism "adopts sub symbolic neural based mechanism and permits a biologically realistic explanation of mind that links to brain functions" [34].

Wang et al (2003) questions whether self-awareness and consciousness can be implemented computationally. Also they argue that a multilevel approach would be much better than a hybrid approach to that integrates the two approaches (connectionism and ACT-R) would bring us closer to the theory of mind. Wang argues though that since both approaches failed in some criteria, hybridizing them might not be a solution. Instead what is suggested is a multi-layered approach to recognize the multilevel nature of the human brain [34].

The Newell test similar to the Turing test has two aspects:
1. Positivistic aspect, which means the theory, has to allow all areas of cognition.
2. Falsifiability aspect, which means that false models have to be restricted and disallowed.

According to Taatgen, the Newell test does not discuss the aspect of Falsifiability in depth which is considered as important. While a Turing test is capable of predicting a conceivable cognitive phenomenon, the Newell test has to have some criteria strengthened to stress the importance of Falsifiability aspect (learning, knowledge integration, use natural language, flexible behavior, and real-time performance) [33].

Intentional strategy or intentional stance includes treating the object whose behavior one is trying to predict as “a rational agent” with beliefs and desires and mental states that exhibit intentionality. Languages enable us to formulate desires. To get what one wants, one has to say what he wants [14]. This is a reference to Schank's machine, which has the aim of a program to simulate human ability to understand stories. The ability to understand a story by human beings allow us to answer questions about the story. Advocates of strong AI, which is unsupported by Schank's work, claim that in question and answer sequence, a machine that is simulating human skills can understand the story and questions about it.

The Chinese room argument is mentioned, where a human is in a room and given a batch of Chinese writing. A second batch is given with a set of rules to correlate the second batch with the first then a third one is given to correlate the third script with the second and first. The first batch is called
“script,” the second is called “story” and the third is called “questions.” Let us use the two claims above to discuss this example. In relation to the first claim, a person who doesn’t know Chinese won’t understand a single word. There are only inputs and outputs that are “indistinguishable.” As for the second claim, the computer and its program do not provide “sufficient conditions of understanding.” One claim by the strong AI is that a story understood in English is exactly the same as manipulating Chinese symbols. Therefore the process is nothing but symbol manipulation where there is a distinction in the case of English that are understood, from Chinese that is not understood [14].

The mind-body relation was first published in two papers: “The Mental and the Physical” by Herbert Feigl in 1958 and “Sensations and Brain Processes” by J.J.C. Smart in 1959. The identity theory introduced the idea that mental events are nothing but brain processes, which was short-lived [39]. Modern conceptions of the mind discuss Descarte’s theory of soul substance and consequent mind-body dualism. He acknowledges the importance of inner awareness and the first-person experience. Secondly, he defined a person as a thinking thing that has doubts, willing, unwilling, and has sensory perception. Ideas are given to humans and based on them perceptions are made. Mistakes are made only when we try to match them to the outside world. Therefore, ideas become a veil of perception between minds and things [40].

2.4 Situation Awareness

The concept of Situation Awareness can be traced to one of Clausewitz’s four main “frictions” of war [9]. However, it is widely applicable beyond the battlefield; in fact, humans in any scenario exhibiting uncertainty about high-stakes, short-term variation lends will benefit from greater situational awareness. Consequently, it has primarily been research in aviation and air traffic control that has expanded our understanding of the field. It has since rapidly spread to other domains due to development in technology[6]. While there are many definitions of Situation Awareness (SA), some are domain specific and some are general, Mica Endsley’s (1988) is the most cited:

“The perception of the elements in the environment within a volume of space and time, the comprehension of their meaning, the projection of their status into the near future, and the prediction of how various actions will affect the fulfillment of one’s goals”[5].

Endsley describes a process known as Situation Assessment [2]. This is a process of information acquisition and interpretation, that achieves Situation Awareness through four critical and integrated stages [4]:

1. Perception: acquiring the available facts.
2. Comprehension: understanding the facts in relation to our own knowledge of such situations.
3. Projection: envisioning how the situation is likely to develop in the future, provided it is not acted upon by any outside force.
4. Prediction: evaluating how outside forces may act upon the situation to affect our projections.

Vidulich et al (1994) describe, “situation awareness is the continuous extraction of environmental information, the integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events”[3].

As an outcome of Situation Assessment, Situation Awareness can be viewed as a product with four levels. Perception, comprehension and projection are described as: level 1 SA, level 2 SA, and level 3 SA, respectively. Jones and Endsley (1996) found that 76% of SA errors in pilots could be traced to problems in perception of needed information (due to either failures or shortcomings in the system or problems with cognitive processes)[10]. While 20% of SA errors were found to involve problems with level 2 SA, and 4% involve problems with level 3 SA. The results of these studies, which demonstrate
a rough analytical approach to the study and structure of intent and perception, are encouraging. However, they underscore the need for a finer-grained understanding of this phenomenon. What researchers such as Endsley and Jones have done is to demonstrate sufficient empirical correlation to show that these phenomena warrant further, more in-depth study. In particular, as systems with increasingly better predictive analysis abilities evolve, their introduction and the subsequent change (or lack thereof) in problems of situational awareness will tell us whether a more systemic or cognitive approach is needed; that is, whether the human brain is a bottleneck in situation awareness improvements.

Situation Assessment is an intensive mental procedure, driven by multiple layers of cognitive processes. It maps the physical distribution of objects and time onto a mental one, and then projects it onto a shared high tempo-physical space. The success of this procedure is contingent on perception and what drives events and actions – intentions [11]. While most SA models place great emphasis on the dynamics between situation and cognitive processes, understanding intentions as drivers of human perception and action is given less priority [6] [2] [11].

Attention, as a cognitive process, has been proven to play a significant role in Situation Awareness [7] [8][12]. Up to 35% of all SA errors were associated with lack of attention; “all needed information was present, but not attended by the operator”[10]. Attention, though central to perception, is determined by a hierarchy of intentions and goals. “[A]gents deployed their attention in ways that are consistent with [their] operational goals”[6]. Goals play critical roles in cognitive models in various capacities: a) “active goals direct the selection of the mental model,” b) “goal and its associated mental model are used to direct attention in selecting information from the environment,” and c) “goals and their associated mental models are used to interpret and integrate the information”[4].

These propositions echo a classical view of intentions as mental states that actively “evolve” into hierarchical forms as a result of acting in the environment. Bruner, among others, argued that goal-directed planning occurs below the threshold of consciousness, and that the components of the resulting action are organized in a more hierarchical fashion [29]. There are different levels of intention with varying relationships to consciousness: high-level, middle-level and low-level intentions. “A hierarchy is formed by these three levels of intention that give us a relation between the means and ends”[30]. High-level intentions are beliefs, emotions, and behaviors that someone wants to cause to another. While middle-level intentions are directed towards goals that are means to achieve high-level intentions, and respectively low-level intentions are the means to achieve both middle- and high-level intentions [30]. The hierarchical view of intentions as they approach action and interaction in the environment is a form of necessary linearization that simplifies relating time in spatio-temporal reality to states and actions.

Human activities can be described, in simple terms, as combined series of proactive and reactive states and actions. That is, factors internal and external to the individual’s mind alike can influence them. Actors have internal states and perceive external states as dictated by the interactive environment. As a system, a human actor operates in a greater environment that can be fairly described as a System of Systems, in which humans are intentional– they have “intentional states” and exhibit appropriate patterns of consistently rational behavior in the sense discussed by Dennett [15] [16]. Rationality here means “acting as best to satisfy [one’s] goals overall, given what [one’s] know and can tell about [one’s] situation”[14]. In this sense “rationality is the mother of intention”[15]. According to this view, intention is the relationship between an object (primarily physical such as an artifact) and an intentional state - a cognitive internal state (e.g. desire, belief, goal, and purpose). “Intentionality is the character of

---

1 Intentionality throughout this paper is considered as the quality of having intentions and should not be confused with a quality of actions or other distinctions mentioned in the literature.
one thing being “of” or “about” something else, for instance by representing it, describing it, referring to it, aiming at it, and so on”[14].

Intention is therefore the relationship between a state and its object or objects. Dennett’s term, intentional stance, is an abstract tool that simplifies analyzing and predicting the behavior of complex systems such as humans. However, it falls short of understanding internal structures and operations [18]. The existence of an intentional stance in a system leads to the next step – identifying the components of a system’s state. There are many possible mental states that can be chosen to characterize a system, such as beliefs, goals, desires, intentions, commitments, fears, hopes, and obligations. An important categorization is to consider states as “attitudes” and group them. One such grouping is as follows:

1. Information attitudes (e.g. knowledge and belief) attitudes an agent has towards information about its environment.
2. Pro attitudes (e.g. goals, desires, and intentions) attitudes an agent has that tend to lead it to perform actions.
3. Normative attitudes (e.g. obligations, permissions and authorization). For more discussion and references on attitudes, see [18].

Most research agendas concentrate primarily on one of these attitudes and investigate relationships among others. Desire-Belief-Intention (DBI) research investigates relationships among beliefs, desires, and intentions. Based on the theory of human practical reasoning DBI theories assume that systems have internal desired states they seek to achieve [17]. By observing the environment and its frequent patterns, systems create beliefs and intentions. They then act towards their desires based on their beliefs and according to their intentions. In simple terms, beliefs represent temporal rules and conditions, while intentions represent procedures or plans aimed to the achievement of desired states.

While DBI theories and the classification of attitudes offer an easy way to speak of systems in abstract terms, an alternative abstract view relying on more fundamental conditions for interacting systems is possible. According to this abstract view, DBI theories can be regarded as special cases. The following is a basic argument that leads to such an abstraction.

If acting in time and space requires making choices, then it is sufficient for the agent to retain what is necessary to make choices. Observing a consistent pattern of rational behavior that demonstrates the same cognitive states and capacities repeatedly and in various combinations justifies assuming that these states and capabilities are the ones necessary to retain [14].

Intentional systems have a capacity to: first, classify non-frequent conditions, in many respects, as frequent by ignoring conditions that are irrelevant to their intentional state and its objects. Thus, it is a capacity to interpret and create multidimensional “states.” The multidimensionality comes from the applied scale of specificity of included conditions and excluded ones. That is, at least theoretically speaking, how the problems of scaling and context are resolved. The second capacity is to make choices, not only on the basis of current “states” but also the pattern of “states” over time.

Therefore a useful abstraction that facilitates general description of systems - beyond the specificity of attitudes (e.g. beliefs and desires) to include systems other than people (e.g. organizations and computers) – is to consider multiple levels based on the following qualities:

1. Types of intentional states, if they exist, defined according to a classification of relations between intentional states and their objects. Temporal, spatial, and ordinal are examples of possible relations. Desires, beliefs and intentions are all intentional states exhibiting different relations with their objects. In this sense, DBI theories are special cases.
2. Types of choices, defined according to a classification of interpretation, justification,
and functionality. For example, interpretable and (or) justifiable choices by primary or (and) secondary set of observers or agents qualify for classifying choices. The mechanism (functionality) of making choices – decision making and its mechanisms, is another possible way to classify choices. DBI theories are special cases, according to this view, since they are mainly based on different flavors of Modal Logic for making choices. All formal techniques of reasoning that human agents utilize are prime examples of rationality types to make interpretable and justifiable choices. The primary set of observers, in these cases, is other human agents, often a targeted audience. Deduction and inductive symbolic logic, like modal logic, are but known examples. Human agents utilize dynamic reasoning types ranging from intuition to deduction, and have the ability to exchange, replicate, mutate and communicate these types.

Specifying two schemes (classes) – one to define intentional states types, and the second to define choices types - allows for a general view of systems based mainly on agency, or interaction with the external environment. Intention schemes, or I-schemes, and choice schemes, also referred to as rationality schemes, or R-schemes, can be used to study systems in a unified way. Fixing two schemes and comparing systems based on what I-types and R-types used is one method of studying all systems. Another approach is to first alter I-types and R-types and consider varying schemes that encompass them, and to then compare results.

All systems may thus be viewed as instances with varying levels of I-schemes and R-schemes. Animals as systems are in this sense equipped with rationality and intentionality that is “different” from human agents. Computers and more general information systems possess degrees of rationality and intentionality schemes. The operating environment is a system of systems where: a) internally, each system has changing degrees of “rationality” and “intentionality” that guide its behavior, and b) externally, each system has a position in space and time that can be considered as an initial state AND a non-initial state.

2.5 Intention Awareness

States and goals are central in Control Theory approach to system behavior analysis. Feedback control mechanisms for example, are designed to reduce the difference between an observed state and a goal - exhibiting a form of efficiency or intelligence. Creating the foundations of Cybernetics, Feedback Control Theory matured during and after World War II, and was recognized as relevant to intelligent behavior [22] [19] [20]. Both feedback and feed-forward control cycles assume the primacy of goals. The difference lies in the type and methodology of the applied rationality; i.e. in selecting strategies. “Feedback control reacts to sensory observations with actions designed to correct errors between observed and desired states,” while “feed-forward uses a model of the world to predict what actions are required to achieve the goal”[23]. Feedback control produces reactive behavior, and is inherently data-driven. Observers can gather some nonzero amount of data about events that have occurred in the past, but since such hard information is unavailable when considering future events, beliefs and perceptions must be shaped by some sort of projection, or set of ideas about data, than about data itself. In this sense, feed forward control produces anticipatory (proactive) behavior.

Control or regulation is seen as a reduction to variety caused by the operating environment. According to Ashby’s Law of Requisite Variety, “only variety in [response] can force down the variety due to [disturbance]; variety can destroy variety”[21]. In other words, if a system is to react to the environment’s variety, this system’s variety has to be greater or equal to that of the environment, in order to seek its goals. Reactive response to potentially unlimited variety creates a problem that requires maximizing the system’s internal variety so as to be prepared for any expected or unexpected
contingency. A reactive system with goals, applying only feedback control, will evolve to also apply feed-forward control, so long as its own existence is also a goal. A feed-forward control assumes that a system models the environment and has the capacity to run that model backwards to predict states and propose actions. These conclusions are consistent with Ashby’s stronger version of the Law of Requisite Variety: “Any regulator able to confine the fluctuations in the system to be regulated must not only have adequate amounts of variety available to control that system but also be or have a homomorphic representation of that system” [24].

These views are consistent with Maturana and Varela’s theory of Autopoiesis, developed in the early seventies [26][27]. Autopoiesis drew from concepts of self-referentiality while closure developed in the second cybernetics, especially by von Forster [26][27]. The theory covers both the internal aspect of living systems’ basic organization (metabolism) and its external aspect (agency). Autopoiesis presents the key notions of: 1) operational closure, 2) component production network (abstraction of the idea of metabolism), and 3) spatio-topological unity (individuality and physical border). The theory’s central claim posits that closure or recursivity in organization is generated by by means of the components and production processes of the system that becomes intertwined in the context of a complementary relationship between a network and a physical border (which are, at the same time, condition for and result of each other) [11][27]. System identity appears when a coherent set of couplings linking components, processes, and flows of energy is established [26]. The system emerges from deep intertwining between constructive and energetic relations, between organizational structure and the specific materiality of its components, which explains the holistic character of its component [28]. A system develops the capacity to manage its internal flow of matter and energy so that it can simultaneously regulate, modify, and control internal self-constructive processes and processes of exchange with the environment. The system must therefore be able to generate and regenerate all constraints — including part of its boundary conditions — that define it as such, together with its own particular way of interacting with the environment.

A summary of results can be drawn from the above discussion:

1. Self existence, as a goal for an intentional system, is necessary for the system to constantly increase internal variety, given a constantly changing operating environment.
2. If achieving an intention is contingent on regulating the operating environment, especially the physical distribution of objects, then the Rule of Requisite Variety has to be satisfied at least once in a sequence of states - system’s states and environment’ states - leading to that intention.
3. From 1 and 2, intentions are hierarchical or always prioritized by rationality. In more general terms, they are temporal order relations that specify a priority of intentions. A system will exhibit interesting forms of behavior when it starts forming its own intentions and seeks to satisfy them. The latter is consistent with a claim regarding Artificial Intelligence (AI) and intentionality noted in [18] citing [25] “that Sandy[,] of the Coffee shop Conversation[,] claims that the really interesting things in AI will only begin to happen when the program itself adopts the intentional stance towards itself.”
4. Intentional systems that reach their goals in a constantly changing operating environment have higher overall variety than systems that operate in unchanging environments.
5. Systems’ internal states – both intentionality and rationality – are richer than their external states – physical positions and redistribution.
6. Every resultant state of the operating environment – a system of systems – is a manifestation of competing rationality, intentionality, and spatio-temporality of individual systems. The notion of competition explain how potential internal variety of individual systems is greater than the actual states the operating system can exhibit, unless
intentionality and spatio-temporality are synchronized and coordinated among all relevant individual systems. The latter means that individual systems perceive time as infinite and thus negate individual systems’ need for self-existence. Further, competition can be classified in varying degrees ranging from cooperation to conflict.

7. Individual system intentionality can be independent, shared, complementary, and conflicting.

8. Because intentionality is hierarchical and precedes its spatio-temporality, then there are equivocal mappings between an action and its intentions. There are equivocal mapping between the physical state of a system and its mental states, i.e. its rationality and intentionality.

Brunner, among other scholars, mentions that goal-directed planning happens below consciousness and that components of the resulting action are organized in a more hierarchical fashion [29] [30]. There are different levels of intentions with a different relationship to consciousness. High level intentions refer to beliefs, emotions, and behavior that someone wants to cause to another. Middle Level intentions are directed towards goals that are means to achieve high-level intentions. Low level intentions are directed towards “the means of bringing about the middle level intentions and the high-level intention in turn”[30]. A hierarchy is formed by these three levels of intention that give us a relation between the means and ends. To understand a person’s intents, one has to infer the other individuals’ high level intentions from the middle and low level intentions. This presupposes a nuanced understanding of the environment in which those other individuals’ intentions are formed in addition to relevant information about the individuals themselves.

Lattice theory can help us shed some light on this complex problem. For example, the law of modularity tells us that for any three elements of a lattice (here termed a, b, and c),

\[ x \leq b \rightarrow x \lor (a \land b) = (x \lor a) \land b, \text{ with AND operation representing the join operation, and the OR operation representing the meet operations} \]

[40]. What this means is that in a modular lattice, the greatest lower bound on a partial order within a given lattice is inversely isomorphic with the least upper bound, and vice versa. We can apply this concept to the intentional hierarchy by arguing that the construction of intentions demands an isomorphism between them, because they necessarily influence each other in some way. For instance, low- and middle-level intentions are brought about by the development of higher level intentions, and the latter cannot be expressed or achieved without some combination of the former. Interactions among the two types of lower-level intentions are similarly isomorphic, as are the relationships between high-level intentions. They all exist within the same set of physical parameters, so some subset intersection, indicating conflict between high-level intentions, will influence change or unexpected results at lower levels. Representing intentions as elements in modular lattices will allow us to better understand the structure of the intentional environment, especially with regards to the interactions among them. We can represent this relationship as between sets of intentions: let the set of all of an individual’s intentions be S, where \( \{S: (l \cup m), h\} \), and l is the set of low-level intentions, m is the set of mid-level intentions, and h is the set of high-level intentions. The hierarchical relationship between the different intention sets can be described as follows: Assuming \( \{l, m, h\} \subseteq S, \forall i \in (l \cup m), \exists j \in h \) such that a mutually isomorphic relationship can be defined between i and j.

Correctly prioritizing information in a dynamic environment remains a challenge aspect of SA; situation awareness is not merely the ability to acquire mass amounts of information; doing so often results in information overload [31]. Instead of the sheer amount of information, the key distinction is its usefulness. A situationally aware system (human or otherwise) must be able to determine with minimal temporal cost what information is useful and what is not, with enough accuracy to produce successful results. A rudimentary example
of this requirement would be to compare three e-mail servers: one has no spam filter, another has a spam filter that occasionally marks an important e-mail as spam (preventing the user from seeing it), and yet another has a filter that occasionally lets a spam email in. The latter clearly exhibits superior situational awareness because the design of its algorithm achieves the goal of preventing inundation with spam but does not sacrifice the superior goal of conveying important information. It is thus built with awareness of the users’ and designers’ intent.

3 Conclusion

Situational awareness embodies a wide set of informational and analytical requirements. To employ it properly, a system must be able to determine the appropriate level of focus for global, or system-level information input, and input at the local level as well, combining these into a useful if not comprehensive situational picture. This requires both goal-driven and data-driven processing. The former is directed across the environment in accordance with currently unresolved goals, whereas the latter receives environmental cues (information about events that have taken place in the environment) and decides whether there needs to be new active goals to properly align with intent. Dynamic switching between the two processing models is important for successful performance in many environments. A popular view of SA as purposeful behavior that is directed towards achieving a goal in a specific task environment has been readily adopted by military and warfare analysts, and holds that because situations are by nature dynamic and subject to rapid, unexpected variation, situation awareness must constantly adapt, lest it become outdated and consequently useless [6]. Because commander’s intent drives the mission, regardless of operational size or scope, it is tempting to cite military science as the primary beneficiary of advances in situational and intentional awareness systems [37]. However, expansion upon some of the notions that drive the process of military organization and action reveals that we can apply the same concepts to any system requiring situational awareness, including cyber-security, aviation and myriad other technologically advanced fields.

Acknowledgments

The author would like to thank Ian Andrews, Sergey Kanareykin, David Johnson, Ammar Qusaibaty and Adam Howard for their contribution to the ideas discussed in this paper and their assistance in reviewing and refining this article.

Author Biography

Newton Howard (nhmit@mit.edu) is the director of the Mind Machine Project and a resident scientist at the Massachusetts Institute of Technology (MIT). He received his Doctoral degree in Cognitive Informatics and Mathematics from La Sorbonne, France where he was also awarded the Habilitation a Diriger des Recherches for his leading work on the Physics of Cognition (PoC) and its applications to complex medical, economical and security equilibriums. While a graduate member of the Faculty of Mathematical Sciences at the University of Oxford, England, he proposed the Theory of Intention Awareness (IA), which made a significant impact on the design of command and control systems and information exchange systems at tactical operational and strategic levels. He has served as the Chairman of the Center for Advanced Defense Studies (CADS), the leading Washington, D.C, National Security Group and is currently its board director. He is a national security advisor to several U.S. Government organizations. He recently published The Mood State Indicators (MSI) algorithm, which modeled and explained the mental processes involved in human speech and writing, to predict emotional states. His natural-language approaches to systems understanding and design pointed to the multi-dimensional structures embedded and nested within speech-based cognitive systems, which have led to much advancement in building more accurate cognitive engines for
modeling behavioral and cognitive feedback. Dr. Howard works with multi-disciplinary teams of physicists, chemists, biologists, brain scientists, computer scientists, and engineers to reach a deeper understanding of the brain. Dr. Howard’s research efforts aim to improve the quality of life for so many who suffer from degenerating conditions currently considered incurable. Advancing the field of brain sciences opens new opportunities for solving brain disorders and finding new means for developing artificial intelligence. Dr. Howard’s most recent work focuses on the development of functional brain and neuron interfacing abilities. To better understand the structure and character of this information transfer he concentrated on theoretical mathematical models to represent the exchange of information inside the human brain. This work has proven applicable in the diagnosis and study of brain disorders and has aided in developing and implementing necessary pharmacological and therapeutic tools for physicians. He has also developed individualized strategies to incorporate solutions for psychiatric and brain prosthetics. Through collaborative research efforts with MIT and Boston University, Howard has studied principles of behavioral and molecular mechanisms affected by brain diseases.

References


33. Taatgen, N.A. (2003). Poppering the Newell Test, Department of Artificial Intelligence, University of Groningen.


37. U.S. Army Field Manual 100-5


BrainSpace: RELATING NEUROSCIENCE TO KNOWLEDGE ABOUT EVERYDAY LIFE

Newton Howard\(^1\) and Henry Lieberman\(^2\)

Mind-Machine Project and Media Laboratory
Massachusetts Institute of Technology
Cambridge, MA 02139

\(^1\)Newton.howard@brainsciences.org
\(^2\)lieber@media.mit.edu

Abstract
BrainSpace is a new way of organizing concepts in brain science and neuroscience. It uses some novel AI techniques to collect knowledge from natural language sources, and novel inference techniques to semantically organize this knowledge and reason about it. Knowledge about the brain and the human nervous system ultimately relates to human thought, perception, and behavior. People use natural languages like English to talk about commonsense concepts, but the brain processes that produce them are described in a highly technical vocabulary. BrainSpace tries to relate technical concepts in anatomy and chemistry to a general body of commonsense knowledge, Open Mind Common Sense. It uses the novel inference technique of Blending, to perform joint inference between technical neuroscience knowledge and Commonsense knowledge, and vice versa. This can permit reasoning from specific experimental results and concepts to descriptions of everyday life, and back again to other relevant experimental concepts and data. We hope it will, among other things, yield new insights into the impact of specific stimuli, drugs, surgery or other interventions, on behavior and perception. The current paper reports preliminary results showing that BrainSpace’s inference organizes concepts in an intuitively plausible manner, indicating that it can serve as a foundation for interpreting more specific experimental and medical data.

1 Introduction

The human brain is one of the most complex functioning objects in existence, consisting of more than 100 billion neurons, and countless other parts. Knowledge about the brain and human nervous system, however, comes relatively slowly, through painstaking experiments. Each experiment addresses only a tiny fraction of the brain’s integrated processes. Our knowledge of the brain’s basic operating principles, structure, and dynamics is still relatively sparse and uncertain. It is important to be able to causally connect the specifics of brain anatomy and chemistry to the specifics of human perception and motor control. Better tools for integrating, correlating and interrelating the wealth of knowledge that does exist about the brain, and relating it to human perception and action, holds the promise of contributing greatly to our knowledge of neuroscience.
2 AI Tools for Commonsense Knowledge Can Help Understand Brain Structure and Function

Recently, fundamental advances have been made in Artificial Intelligence in representing and reasoning with knowledge about people's everyday experience, which is referred to as Commonsense knowledge. Commonsense knowledge can often be expressed in simple sentences like “Apples are red” or “Glasses help you see.” Working with Commonsense knowledge in AI has some differences from working with more factual knowledge, like “Barack Obama is President of the United States.” Commonsense knowledge can often be vague, subject to exceptions, influenced by culture and context, etc. Our Commonsense knowledge base, Open Mind Common Sense (OMCS) (Singh et al. 2002) contains over a million assertions in English, which we estimate represents 1% of an average person's Commonsense knowledge.

The traditional way of working with factual knowledge is to represent it using ontologies (hierarchies of precise definitions of terms), and assertions (statements asserted as true, expressed in the vocabulary of the ontology). Inference rules from mathematical logic take existing true assertions and generate new true assertions or confirm or deny the truth of a candidate assertion. This inference can be very precise, but it doesn't work well with Commonsense knowledge, and the computation of inference can easily spiral out of control.

A new approach to reasoning with Commonsense knowledge is AnalogySpace [2][3]. Instead of insisting on a fixed hierarchical ontology, it creates concepts from parsing natural language statements of Commonsense knowledge, allowing noun phrases and verb phrases as well as individual words to be concepts. Redundancy and contradiction are permitted. A matrix of concepts vs. features (an assertion minus a concept) is created, so the cells of the matrix represent all possible assertions about the concepts and features being studied. Assertions that actually appear in the knowledge base fill in the corresponding cell in the matrix with a value indicating confidence in that assertion, if known. This matrix is usually very sparse, since any given knowledge base only contains a small fraction of possible assertions.

Reasoning is performed by reducing the dimensionality of that space, using the mathematical techniques of Principal Component Analysis. This has the effect of “filling in” the missing cells in the matrix. It organizes the space according to its most significant patterns, which can often correspond to semantically meaningful distinctions such as “good vs. bad” or “easy vs. hard.” The technique is more tolerant of noise, imprecision, contradiction, context and point of view, than traditional logical inference.

A related technique, blending (Havasi et al. 2009), can perform cross-domain inference (Arnold et al. 2010) reasoning jointly across more than one independently created knowledge base, even if the knowledge bases may not initially completely agree on their vocabulary, or may not completely agree on certain assertions [3][4].

Blending is often useful in reasoning jointly between a specialized, factual knowledge base, and a more general, Commonsense knowledge base. One example of such an application, in game programming for novice programmers, is reported in Arnold and Lieberman 2010. There, knowledge from a program library about which code performs which functions is combined with knowledge about the domain of the program (interactive video games), and with general Commonsense as background knowledge [4].

BrainSpace is a project to use these new advances in reasoning to apply to knowledge in the neuroscience area. It performs cross-domain inference between specific neuroscience knowledge and general Commonsense knowledge. It comprises a number of efforts in knowledge acquisition in the neuroscience area.

3 Uses of BrainSpace

BrainSpace can have a wide variety of applications in organizing information about neuroscience.

- Using relations between human activities to infer relations between experimental results
Many neuroscience experiments whose goal is to better understand the functional organization of the brain involve asking participants to observe some stimuli or perform some action, while the experimenters take some physiological measurements, such as fMRI. They see if they can correlate specific real-world activities with activity of a particular brain region or chemical process. This permits them to reason about the role of that brain area or chemical pathway in causing the human activity.

But any such individual experiment must necessarily focus on a very narrow slice of human behavior. Reasoning about the human behavior can expand the range of behavior for which the physiological observation may have relevance. For example, if an experimental subject is told to push a button when they see a picture of an apple, one might conclude that the brain activity observed in an fMRI might be involved in the visual recognition of the apple. But it is difficult to make good conclusions about which aspect of the apple stimulus actually caused the physiological response. If certain aspects of the stimulus are identified as experimental variables, it might be possible to determine if these particular aspects caused the response, but each experiment can only be controlled for a few variables.

It may happen that another, unrelated experiment might also present stimuli that may be related in a Commonsense way to the first experiment. For example, a subject may be told to push a button when seeing another object, such as a fire hydrant. If that activated brain regions in a similar way, we might conclude that perception of the color red was the salient element. But it requires reasoning that the relation that an apple and a fire hydrant are most related by their color. This would much more strongly establish the relation between the brain region and color perception. If, on the other hand, stimuli about fruits or other foods were related, we might hypothesize that a “mouth-watering” kind of reaction was taking place. In practice, we would expect BrainSpace reasoning to hypothesize much less obvious connections, which could then be evaluated by targeted experiments.

- **Organizing the space of experimental results for intelligent search and navigation**

Rather than think about individual experiments and results, as in the preceding paragraph, we could also use AnalogySpace inference and blending to perform large-scale organization of scientific results. The scientific literature is growing at an enormous rate, and searching via keywords has its limits. Insisting that everything be classified by strict ontologies poses problems for users, who may not know the correct terms in advance.

BrainSpace could enable researchers to ask general questions about “what’s relevant” to a given concept. They could navigate through dimensions that represent relevance or interestingness computed dynamically and specified by example. A researcher searching for “related” experiments or a doctor searching for “related” cases could better identify which knowledge was related.

- **Understanding the effects of abnormal conditions, interventions like drugs or surgery, etc.**

One can recognize an unusual condition by contrast to the usual case. How do we know what is usual? That’s Commonsense. We have had some experience in recognizing unusual or outlying situations in other domains by comparing them to the Commonsense model of the usual or normal condition [5]. In other projects, we have used Commonsense to provide semi-automated assistance for diagnostic procedures, by systematic comparison of unusual-to-usual conditions under differing levels of detail. A “Sherlock Holmes” investigative approach tries to see whether high-level descriptions of a situation meet Commonsense expectations, and where they diverge becomes a signal to look at the situation at a more detailed level.

This is not dissimilar to diagnostic procedures used by physicians. A Commonsense model of the brain can be useful in diagnosing and understanding the effects of abnormal or unusual situations, the effect of diseases or injury on the system, the effect of drugs or surgery or other interventions. It can
also suggest where missing information might yield essential clues, so that diagnostic tests might be ordered to collect information that would permit a specific inference.

**Collecting and representing basic neuroscience concepts**

In the summer of 2010, Elizabeth Hawkins initiated the knowledge acquisition activities by making a collection of preliminary assertions about basic concepts in neuroscience, as a foundation for BrainSpace. Open Mind Commonsense constructs a semantic network from assertions in natural language by using a variety of linguistic patterns. It has about 20 native relations that generally express relationships between concepts, such as “X is a kind of Y” or “You use X to Y.” The relation set was extended with a few relations specifically for the neuroscience domain, e.g. “Connected To,” “Activates,” “Regulates” and “Inhibits.”

We also intend to run experiments that encourage our neuroscientist and medical expert collaborators to play short sessions of simple games with a purpose that collect knowledge. Our experience with such games, such as Common Consensus (Lieberman et al. 2007) shows that we can collect knowledge from experts at rates approaching 100 assertions/participant/hour [6].

The most basic concepts, e.g. “neuron,” were started from introductory articles in Wikipedia, since it had the simplest descriptions, intended to be understood by a layperson. The next stage used as a source a general-interest book on the brain, *The Human Brain Book* (Carter 2009). Finally, a medical-school-level textbook, *Principles of Neural Science* (Kandel et al. 2000) was used, so that we had descriptions in increasing order of technical sophistication.

Here are some examples of BrainSpace assertions:

- A neuron is a cell.
- A neuron is part of the nervous system.
- A neuron has an axon.

A motor neuron causes muscle contraction.
A dendrite can be found near a synapse.
An ion channel is chemically gated.
GABA is a neurotransmitter.
Serotonin causes inhibition.
The brain stem is used for breathing.
The neocortex is used for learning and memory.
A neurotoxin can cause brain damage.
The hypothalamus regulates pituitary secretions.

We should caution that, as Commonsense knowledge, it only represents the simplest level of knowledge, and has no easy way to deal with knowledge that might need to be expressed in multiple sentences, relations between context-dependent assertions, meta-relations, etc. These will be the subject of future research.

There are currently approximately 500 assertions about basic neuroscience concepts in BrainSpace. Experience with collecting knowledge in other domains has led us to expect that a basic knowledge of a field becomes really significant at the level of at least a few thousand assertions, so it is still a bit early to properly evaluate BrainSpace as a knowledge collection. However, we have experienced some encouraging signs that the knowledge base is, in fact, beginning to “make sense” and producing semantically coherent results, detailed below. It will continue to grow.

We have constructed an AnalogySpace from the current BrainSpace assertions, and present some example results below. AnalogySpace creates a multidimensional vector space (typically between 50 and 100 dimensions), where the angle between vectors in the space is an indicator of semantic similarity. We have a visualizer/navigator (Speer et. al. 2010) which projects these dimensions onto the two dimensions of the screen [9]. An interactive “grabbing” operation permits real-time re-orientation of the semantic space.
4 BrainSpace Examples

The following BrainSpace examples are meant to show that AnalogySpace inference on BrainSpace assertions results in a plausible automatic organization of the semantic space. Examples of brain regions appear properly grouped in the semantic space, even though not all of the terms were explicitly identified in the assertions as being regions of the brain. Examples of human behavioral activities were also grouped, even though we did not introduce a specific concept named behavioral activity.

Figure 1 shows a view of BrainSpace from one perspective. In this screen shot, one can see along the vertical axis, the anatomical features of the brain. Things that are controlled by the brain but not a physical part of it appear on the horizontal axis. For instance, motor skills, balance, and breathing are on one end, and social behavior and emotion on the other end. The Y axis in this image is populated with physical features of the brain such as the cerebral cortex.

Figure 2 shows some interesting clusterings on the left of the image. In the lower left, the endocrine system has been grouped with sleep, mood, muscle contraction, mood, and learn, things that might be affected by the endocrine system. In the upper left, movement, posture, head movement, social behavior, and balance are all around conscious action.
The next figure shows various anatomical features that are located in the hindbrain and midbrain, forming a semantic cluster. You can see the pons, the brainstem, the cerebellum, the spinal cord, etc.

Very interestingly, diencephalon is not actually in the hindbrain, but it is in the hindmost part of the forebrain. This shows that, as is often the case with Commonsense knowledge, it can sometimes get things “almost right.” That may not be good enough for situations where very accurate reasoning is required; for that, precise logical reasoning may be required. But for many situations, approximate reasoning may indeed be good enough. In other situations, a hybrid approach could be adopted where Commonsense reasoning is used to generate first-pass hypotheses or to get overall intuition, followed up by more precise reasoning for refinement or verification. People do this all the time. In the upper right of the next image, you can see a number of neurotransmitters all clustering around “neuron.”
Figure 5 actually illustrates a failing of the current version of BrainSpace. In the lower left, you can see breathe and regulate respiration clustering reasonably close but not adjacent. However, if you move one, the other does not move with it, showing that they are not recognized as being related. Another pair of concepts that are intuitively related but were not recognized is taste and eat. However, we fully expect that this is the sort of problem that will be cured immediately when we perform the blend between BrainSpace and ConceptNet’s general Commonsense knowledge, closing this sort of semantic gap. Finally, Figure 6 demonstrates another clustering of brain anatomy along an axis, with things that are caused by the brain’s actions forming their own separate group.
5 Analyzing a Functional Neuroimaging Database

Kevin Gold, Catherine Havasi, Michael Anderson and Kenneth Arnold performed a preliminary analysis of a database containing functional neuroimaging (fMRI) data, connecting activation of brain regions to specific experimental tasks [10]. Using the principal component analysis technique of Singular Value Decomposition (SVD), we obtained some provocative and, in retrospect, intuitively plausible, results. The analysis revealed a primary component that appears to correspond to the difficulty of the task posed to experimental subjects. In relatively difficult tasks, many brain areas contribute, but brain areas of high structural centrality represent a greater contribution.

Analysis with a related mathematical technique, Non-Negative Matrix Factorization (NMF) revealed that particular brain areas dominate the activation patterns of their neighbors, and that areas vary widely in how strongly they predict the activation of their counterparts in the opposing hemisphere. We have not yet integrated this space with Hawkins’ basic neuroscience BrainSpace described earlier, relying instead on correlations inherent to the fMRI database. That remains as future work and should yield additional insights.

6 Integration with Traditional Neuroscience Knowledge Sources

There have also been some attempts to use more traditional machine-learning techniques such as classifiers directly on brain data to learn correlations between neural activity and higher level concepts [11]. It might be possible to integrate the two approaches. Finally, we also plan as future work integration with more traditional sources of neuroscience information. Due in part to motivation provided by the Semantic Web, a movement to represent knowledge in standard formats to encourage sharing, many efforts are underway to codify and standardize neuroscience knowledge [12]. Examples include the Whole Brain Catalog (http://wholebraincatalog.org) and the Neuroscience Information Framework (http://www.neuinfo.org). As we develop further methods of incorporating corpora from these sources, and NLP methods for analyzing natural language descriptions to be found there, we can grow our knowledge base by leaps and bounds.

Acknowledgements

We thank Ed Boyden, Kevin Gold, Mark Marchionni, Marvin Minsky, Sebastian Seung, Elizabeth Hawkins, Catherine Havasi, and Kenneth Arnold for contributing their advice and expertise in neuroscience. We gratefully acknowledge the support of the MIT Mind-Machine Project (MMP).

References


LXIO: The Mood Detection Robopsych

Newton Howard¹ and Mathieu Guidere²

¹ Massachusetts Institute of Technology, Cambridge, MA 02139 USA  nhmit@mit.edu
² University of Toulouse, F-31058, France  mathieu.guidere@univ-tlse2.fr

Abstract
Words and languages are the direct representations of our thoughts and perceptions. Accordingly, the only access to an individual’s thoughts and psyche is through the language that he/she speaks or writes [1]. Psychological disorders and mood states such as depression and PTSD therefore, can be identified and predicted by continuously analyzing a patient’s discourse. In doing so, we utilize the method of predictive linguistics [2], which determines cognitive mood states by computationally modeling the notions of ‘mind axiology’ and ‘emotional states.’ Our proposed system expresses cognitive states in terms of axiology, i.e., the system of positive and negative values associated with concepts that are universally accepted. Axiological notions provide more insights about the various cognitive states because a patient’s words are only intelligible if axiological elements like conception, perception, and intention are taken into consideration. The framework for our analytics engine consists of multiple modules responsible for coherently and systematically retrieving, parsing and processing a patient’s discourse. The backbone of this system relies on a learning algorithm that accounts for various valuation criteria such as time-based, intrinsic, consequent and contextual value analysis, while also building a nested network of mood states and their dependencies. By executing this mechanism and correlating our computational data to results calculated by a psychological assessment testing, we are better positioned to classify a patient’s mood states – e.g., as being depressive or not. This work will have implications in the area of cognition research and practical applications such as automated neurological and psychiatric assessments, advanced human – computer interaction (HCI), as well as others.

Keywords: MDA, predictive linguistics, depression, axiology, LXIO

1 Introduction
Understanding a patient’s discourse by examining their selection of words, their value, and the instances of which they used them in will paint a clear picture of their current state of mind. After all, language is the container of our intent [3] and focusing on such attributes will allow real-time mood analysis that can prove effective and essential in treating disorders. This psycholinguistic analysis is to link verbal expressions to psycho-cognitive states. Analyzing a person’s words (both oral and written) highlights the underlying positive or negative language values over time and/or space [7].

---
¹ Language axiological input / output system for detection of mood states towards the design of a robot psychiatrist
² Patent application number: 61/322158
³ Notice: This publication has a patent pending.
Our proposed system deciphers language in written formats (speech to text is also applicable) and passes its data to a Parser which breaks down each sentence into multiple words and creates a data structure for these particular phrases within specific time frames. These structures are generated based on the properties and characteristics of each word. They grow in capacity and connectivity depending on the relativeness between previous and currently processed data. After the features of each word are extracted, they are then passed through a processing framework that includes multiple independent modules based on time-oriented, axiological, contextual, and consequent criteria. Initially, each word is exposed to the time-oriented engine that dynamically looks up the tenses for each word and logs its results. This iteration is followed by a direct comparison function that fetches the values for relevant words within the Mind State Axiology.

Accordingly, each sentence is then passed into the consequent and contextual modules and their results are logged for further examination. Additionally, this framework serves as a sandbox where various combinations of these criteria are examined and new outputs are generated. The output of this system defines a modified object for the sentence values, which can then be further analyzed by a learning algorithm within various proximities to project a mental state that most likely presents a patient’s mood and is verified by practicing physicians.

### 2 Experimental Paradigm and State Modulations

#### 2.1 Patients

A wide range of patients participated in a study conducted at the University of Geneva where axiology-based values were extracted based on a patient’s mood and feelings regarding different words and sentence phrases. These lists of words were presented to each patient and accordingly their value (+ or -) was then individually retrieved based on each patient’s perception. Common word values were then identified from the wide variety of collected data and then stored in a database referred to as mind default axiology database [2].

#### 2.2 Mind Default Axiology (MDA)

A database of axiology words with multiple dimensions that are based on the frequency of the word-value used along with the cultural/traditional usage of the word and not only its intrinsic value. Using this database as a main lookup source helps assign the initial state value for each word, which has significant effect on the projected value of a given statement.

From a mathematical standpoint, axiological value defines truth within a general cognitive knowledge system, often referred to as axiomatic. Therefore, value is a starting point within a logical system that can be chosen at random. However, the relevance of the axiomatic calculation depends on the relevance of its values and their interpretation [7]. Axiological values are the basis of all formal cognitive processes. Here is a simple calculation that includes values and follows a rule of internal composition (+):

the value 0 exists
All values X are followed by succ(X)
\[ X + 0 = X \]
\[ \text{Succ}(X) + Y = X = \text{succ}(Y) \]

Using these values and defining 1, 2, 3 as \( \text{succ}(0) \), \( \text{succ}(\text{succ}(0)) \), \( \text{succ}(\text{succ}(\text{succ}(0))) \), respectively, then:

\[ \text{succ}(X) = X + 1 \]; and:

\[
\begin{align*}
1 + 2 & = 1 + \text{succ}(1) & \text{Abbreviation extension (2 = succ(1))} \\
1 + 2 & = \text{succ}(1) + 1 & \text{Axiom} \\
1 + 2 & = 2 + \text{succ}(0) & \text{Abbreviation extension (1 = succ(0))} \\
1 + 2 & = 2 + 1 = \text{succ}(2) + 0 = 0 + \text{succ}(2) & \text{Axiom} \\
1 + 2 & = 3 = 0 + 3 & \text{Use of the abbreviation (succ(2) = 3)} \\
0 + 1 & = 1 + 0 = 1 & \text{Axiom} \\
X + \text{succ}(X) & = \text{succ}(X) + X & \text{Axiom}
\end{align*}
\]

All assertions that cannot be inferred from the values and whose negation also cannot be inferred from them can logically be added as axioms [7]. Axiological attribution is based on mathematical logic that attempts to assign a temporal value to language, such that the meaning given to a word is
the consequence of its temporal value in relation to its meaning. The goal is to find the best possible axiological value amongst many possibilities [7].

2.3 State Projections

We can treat the process of generating sentence values in terms of state instances. These states vary based on the various measurement criteria we mentioned previously. With that in mind, a given word might pass through different instances and its value could change depending on the stage level of interaction within a given statement.

3 Framework Analysis

Figure 1 highlights the main structure of our proposed system. For the sake of simplicity, we will not fully expand our description of each segment, but we will present a concise and accurate depiction of functionality and usability. Figure 1 describes the higher-level design of the analyzer. Many parts of this architecture can be freely interchanged or grouped based on testing apparatus conducted by scientists and researchers. This unique feature allows us to adjust and fine-tune our system to compensate for errors resulting from a patient's inaccurate descriptions or words that have a particular and unique sentimental value to the patient being examined.

The system also takes into account various forms of linguistic expressions regardless of the language or dialect it was described in. By inputting the axiological database of specific languages and applying an accurate parser we can account for differences in language and structure.

Initially, our approach was to use a parser that simplifies the representation of each clause and by simply assigning negative values to words presented in the past and positive values to words presented in the future. Consequently, we would have a plain depiction of word's effect on the patient. This technique does not, however, take into account the effect of words upon one another and the context of when these words were used. There are many other factors that contribute to how a word is used, perceived or emphasized, which led us to the following state analyzer architecture that compensates for such criteria.

3.1 Parser Transform

In this segment we used a parser developed at Massachusetts Institute of Technology (MIT) called the START parser. This parser allows us to retrieve the main building blocks of each word use and consequently understand its grammatical structure.

3.2 Pre-processing State

At this state, the server breaks up sequences of sentences and sends back a sequence of clauses associated with each sentence. There are two types of clauses, which are relevant to the computation of word values. These are “has_tense” and “word_root” clauses. Initially, if the verb is not in our MDA (Mind Default Axiology) it will be counted as “+” for present/future tense, and “-” for past tense. In order to count “has_tense” clauses, we leverage the fact that “word_root” clauses always come after “has_tense” clauses. As such, when the server is processing a “has_tense” clause, it first looks up the word in the dictionary. Then, if the word is found, the sentence value is left unchanged (since the value of the word will be appropriated added later when the corresponding “word_root” clause is processed),
and if the word is not found, the sentence value is incremented or decremented according to the tense. In the latter case, the verb is stored to a list variable “tense_counted” so that they can be rendered properly in the front end.

3.3 Modular Engine

This engine consists of several modular constructions joined by logical data dependencies where each plays a significant role in evaluating a given discourse.

The framework is divided into the following:

3.3.1 Dynamic Axiology Module:

The data here is compared to a mind default axiology (MDA) database and a given pre-defined dictionary. The role of this module is to identify and retrieve the value associated with identical words. The processing of “word_root” clauses in this module is simple, if the word is looked up in the database or dictionary, and if its value is found, it is counted, else it is not counted. The value of the word is determined either by the lookup, which has just occurred, or by the previous counting of a “has_tense” clause, which can be found in the list “tense_counted.”

The resultant of such lookup will be stored at the given module. This mechanism will iteratively continue until the whole sentence has been processed. Labeling each word with its value enables us to compute and update sentence values in linear time. From Figure 2, both the MDA and dictionary contribute to the value for each word string.

```c
set axiology_value (set mda, set dictionary)
{
    int i;
    for (i=0; i <= sentence_length; i++) {
        axiology_value = mda ^ dictionary;
    }
    return axiology_value;
}
```

This pseudo computes the axiology value of the word based on AND’ing the values from both the MDA and dictionary. Using this logical operation we can account for the values retrieved from both databases while outputting the desired result. A straightforward string-matching algorithm [5] is implemented to identify matching word patterns.

```c
int findmatch (char *p, char *t) {
    int i, j;    /* counters */
    int m, n;    /* string lengths */
    m = strlen(p);
    n = strlen(t);
    for (i=0; i<=(n-m); i++) {
        j = 0;
        while ((j<m) && (t[i+j] == p[j]))
            j = j+1;
        if (j==m) return (i);
    }
    return (-1);
}
```

3.3.2 Time-Based Module:

This module accounts for time orientation. Time is very essential in analyzing a patient’s mood. Research suggests that individuals will separate their personal experiences, the basis of memories, into the psychological time frames of past, present, and future [4]. These memories, good or bad, have different implications on current behavior and future directions. Psychologist Phil Zimbardo maintains that individuals with a more future-based time perspective, for example adopting goals to achieve, exhibit greater psychological satisfaction than those who do not [4]. Hence, a tense that represents the past is considered to be negative, while a tense representing the future is considered positive. Many sentences carry their own representation of the tense.
as a whole in which case we need to account for such tenses. Therefore, the module examines the tense of a given word and verifies its time projection.

3.3.3 Contextual-Predictive Module:

Based on sentence analysis, this module is responsible for identifying words that can affect the value of succeeding words. Therefore, some words that can be linked to others will affect their values. The module takes into account these combinations and creates a database that identifies them in various languages. Hence, contextual words are either negative or positive. For example, “my life” is positive, “life” is intrinsically positive and “my” is positive since it is linked to it. There are several algorithms to represent this mechanism, perhaps one of which is the use of the nearest-neighbor heuristic. This approach not only allows us to identify the nearest words within a sentence but it can also be used to correlate those words that have a contextual connection within different sentences.

3.3.4 Consequent Module:

This module is used to allocate words that carry a meaning in themselves. Many compound nouns such as ‘nothing,’ ‘nowhere,’ ‘somebody,’ etc. have either a positive or negative value based on what they represent. For example, the word nothing is opposite to something, which has a positive value, hence nothing is negative. After all of this, modules have computed the values for each word, logs are generated per each stage and at each level of a given paragraph. Therefore, if we were to trace two words throughout the whole modular process we can witness the following. For example, the sentence: “my life is good” then “my” and “life” are presented as Figure 3 demonstrates. In this example, word values are computed for each module separately. The value from each modular iteration is then convoluted with the value from the other module.

3.4 Modified State

This is the initial state representing the resultant mood value based on the modular architecture. This value is the approximate representation of each mood state, however it is still dependent on the history of the patient and their different word representations.

3.5 Project Mental State

We are applying a learning algorithm that looks at various instances and tracks a patient’s word analysis history. This algorithm takes into account the resultant modified state and redefines the word values based on the history of the patient. Once every word is clearly defined or calculated, the mood state is represented based on the unary summation of the words forming the full sentence. To further explain this notion it is important to look at how this unary system behaves.

4 Unary System Topology

The unary system can identify mental mood states by analyzing abstract structures of an individual’s mind [7]. Positive or negative mental states are defined by the formation of new concepts, which are in part or aggregately created by various forms of activations (electrical, chemical and biological) throughout the brain [7]. These activations in a series or sequence establish an activation set.
For each activated region in the brain an activation set constitutes a connected framework, defined as a node [7]. A node’s circumference adjusts based on an activated region's duration, and the reflexivity of a node is a result of varying instances of region re-activation. Nodes, representing forms of activation, can be linked to each other. The nodes' connections vary in shape and time orientation, while the actual segment that forms the connection itself represents time orientation [7]. A new activation set is created once this connected structure of nodes is formed. The new activation set can also be connected with other activation sets, forming a concept set. Based on unary system calculation, nodes in the same set are added together in terms of waveform signals that are weighted by a statistical coefficient to produce an resultant active node [7]. A concept set, formed from connected activated sets, generates an axiological value. This axiological value, after being projected on a positive and negative plane, designates a mood state. The activation set with the higher positive value defines the value of the concept set [7]. Hence, the dominant activation set subsides the effect of other activation sets. The value of a concept set ranges on a scale from positive infinity to negative infinity; after projection the value becomes a unitary positive or negative.

For example, if we were to think of “Sky” as a concept set and in order for us to determine what it's axiological value is we need to follow the current procedure. The concept “Sky” is a resultant of various activation sets in which each contribution was based on the level of its activation within a specific time frame and orientation. If we were to consider another concept such as “Dark” most likely the same computation applies but a negative axiological value would be assigned. At another point in time and according to the same patient the concept “Morning” can be formed based on these two concepts and it will result with a positive axiology since a “Dark Sky” might resemble the end of a day and hence announcing the coming of the “Morning”. This axiological value is valid within these time-division constraints and it might change if the constraints were to change, similarly for other patients. Therefore, the use of a learning algorithm will enhance our computational accuracy while reducing evaluated errors for specific patients.

### 5 Unary Algebraic Topological Framework

Below is an introduction to a mathematical framework for Unary Topological Axiologies describing mental states [8]. The structure highlights a continuous signal representation: Let's begin with a set $S$ (uncountably infinite) representing brain regions, which may be activated by some means. We introduce a $\sigma$-algebra $A$ on this set, and call the elements $a \in A$ activation sets (by definition $a \subset S$). Now introduce a second set $W$ whose elements are labeled concepts in the brain, which correspond to words. For some subset $\mathcal{A} \subset A$ there is a mapping $P: a' \in \mathcal{A} \mapsto w \in W$ called the concept activation mapping. The elements $a'$ of $A$ are action potentials. Let $P': w \in W \mapsto a \in \tilde{A}$ be a mapping we call the brain activation mapping. Let $\mu$ be a measure on $S$, and let $F: A \rightarrow \{+,-\}$ be a parity mapping. An axiology is a mapping $\Xi: W \rightarrow \{+,-\}$ generated by computing:

$$f(w) = \int_a F(s) \, d\mu$$

with

$$a = P'(w)$$

and the projecting

$$\Xi(w) = \text{sign}(f)$$

Symbol Definitions [6]:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Brain regions</td>
<td></td>
</tr>
<tr>
<td>$A$</td>
<td>Activation sets</td>
<td>$a \in A \Rightarrow a \subset S$</td>
</tr>
<tr>
<td>$\mathcal{A}$</td>
<td>Concept activation sets</td>
<td>$\mathcal{A} \subset A$</td>
</tr>
<tr>
<td>$W$</td>
<td>Concepts</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>Concept activation mapping</td>
<td>$P: a' \in \mathcal{A} \mapsto w \in W$</td>
</tr>
<tr>
<td>$\Xi$</td>
<td>Axiology</td>
<td>$\Xi: W \rightarrow {+,-}$</td>
</tr>
<tr>
<td>F</td>
<td>Parity mapping</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>μ</td>
<td>Weight mapping</td>
<td></td>
</tr>
</tbody>
</table>

Algebra Image [7]:

Axiology Image [7]:

Conclusion

The cognitive mood analyzer (LXIO) is essential in identifying a patient's state of mind. The fundamental framework for such an analyzer depends highly on axiological values, time-orientation and the inter-relation between the words forming a discourse. This architecture offers a unique interface to how words are represented and the method to evaluate their values. However, some characteristics such as facial and gesture analysis would also increase accuracy and estimation of a patient's mood state. We would further investigate the possibility of using other methods to utilize the experience and knowledge acquired by psychologist's conducting assessment tests. By allowing the clinician's to adjust our modular framework according to their personal analysis of a patient we would hence improve the robustness of our mood projections.

Acknowledgements

We would like to thank Mark M. Atallah, Sergey Kanareykin, and Karina Kabbash for their contribution to the ideas discussed in this paper and their assistance in reviewing and refining this article.

References

Abstract
Several intervention modalities for treating brain dysfunction meet criteria as ideal manipulation devices in that they are non-invasive, anatomically selective and functionally targeted. Of these modalities, ultrasound, defined as cyclic sound pressure with a frequency greater than about 20 kHz, is increasing in its application and popularity as a treatment for brain dysfunction. Therapeutic ultrasound varies in its benefits, reversibility and the types of coupling media that may be used to deliver treatment. The uses of therapeutic ultrasound include lithotripsy, physical therapy, drug delivery, and High Intensity Focused Ultrasound (HIFU) for tissue ablation. In addition, low-intensity low-frequency ultrasound, traditionally used in the treatment of musculoskeletal conditions, is advancing as an experimental treatment modality for brain dysfunction. The popularity of low-intensity low-frequency ultrasound as a treatment for brain dysfunction is likely to increase in response to recent advances in the development of transcranial ultrasound application methods.

Keywords: brain dysfunction, ultrasound, therapeutic ultrasound, low-intensity low-frequency ultrasound, transcranial

1 Introduction
One of the major challenges in the study of brain function and the treatment of brain dysfunction such as depression, attention-deficit-hyperactivity and post-traumatic stress is to find ways to manipulate the brain. These manipulations should be non-invasive, anatomically selective and functionally targeted. Manipulation methods available today include pharmacological interventions, surgical ablation, transcranial electric and magnetic stimulation, and optogenetic techniques, none of which fulfill the criteria for an ideal manipulation device outlined above.

Any non-invasive intervention modality beyond pharmacological treatment requires the transfer of energy across the cranium and to the target brain structures. Transcranial magnetic stimulation (TMS) has been used to transfer energy but its poor spatial resolution and penetration power limit its potential [Hallett, 2000]. The use of ionizing radiation, such as gamma knife radiosurgery [Leksell, 1958;
Jagannathan et al., 2009a, 2009b], offers better spatial resolution and tissue penetration but significant concerns about adverse effects to tissue outside the treatment zones remain. Ultrasound is another modality used for non-invasive energy-transfer. It has been used since the 1950s in the therapy of disorders ranging from kidney stones to musculoskeletal conditions. But until only a few years ago, it was not possible to sufficiently focus the ultrasound beam through the cranium, and therefore this form of therapy lacked precision in the potential treatment of brain disorders. However, recent advances in the development of transcranial focusing methods combined with closed-loop imaging make it now possible to deliver energy through the cranium and to defined structures in the brain. Here, we review, briefly, the defining characteristics of ultrasound and its application in diagnostic imaging before examining current and future applications of therapeutic ultrasound.

2 Ultrasound in Medical Diagnostic Imaging

Ultrasound is defined as cyclic sound pressure with a frequency greater than the upper limit of human hearing (about 20 kHz). Like other sound waves, ultrasound is absorbed and reflected by matter. In the medical setting, ultrasonic reflection signatures are typically used for diagnostic purposes to visualize subcutaneous body structures including tendons, muscles, joints, vessels and internal organs.

Typical diagnostic sonographic scanners operate in the frequency range of 2 to 18 MHz. The choice of frequency is a trade-off between spatial resolution of the image and imaging depth: lower frequencies produce less resolution but image deeper into the body. Higher frequency sound waves have a smaller wavelength and thus are capable of reflecting or scattering from smaller structures. Higher frequency sound waves also have a larger attenuation coefficient and thus are more readily absorbed in tissue, limiting the depth of penetration of the sound wave into the body. The sound wave is partially reflected from the layers between different tissues. Specifically, sound is reflected anywhere there are density changes in the body: e.g. blood cells in blood plasma, small structures in organs, etc.

Diagnostic ultrasonography is generally considered a safe imaging modality [Merritt, 1989; Bioeffects Committee of the American Institute of Ultrasound in Medicine, 2008]. Its use in prenatal diagnostic is well established and most studies have found it to be safe for the foetus [Salvesen, 2007] although there have been reports of ultrasound affecting prenatal neuronal migration in mice [Ang et al., 2006].

3 Ultrasound as a Medical Therapeutic Tool

Introduction

Therapeutic ultrasound may have two types of benefit: thermal effects and non-thermal effects. Thermal effects are due to the absorption of the sound waves. Non-thermal effects are from cavitation, microstreaming and acoustic streaming. Cavitative effects result from the vibration of the tissue causing microscopic air bubbles to form, which transmit the vibrations in a way that directly stimulates cell membranes.

Biophysical Effects

Therapeutic ultrasound is the use of alternating compression and rarefaction of sound waves for therapeutic benefit. When ultrasonic energy is induced into an attenuating material such as tissue, the amplitude of the wave will decrease with distance. This attenuation is due to either the absorption or the scattering of sound waves [O’Brien, 2007]. Ultrasound therapies can be divided into “high” power or “low” power therapies where high power applications include high intensity focused ultrasound (HIFU) and lithotripsy; low power applications include sonophoresis, sonoporation and gene therapy, as used for drug delivery purposes, but it is also used for pain-management and wound-healing purposes [ter Haar, 2007].

Biophysical effects of ultrasound are traditionally separated into 2 types: a thermal effect from absorption and a non-thermal effect from scattering. The absorption of the ultrasonic sound energy leads to tissue heating. The scattering is thought to be that
portion of ultrasonic energy that changes direction and leads to the non-thermal effects [Wilkin et al., 2004].

Thermal effect: When ultrasound travels through tissue a percentage of it is absorbed, leading to the generation of heat within that tissue. The basis of HIFU is to raise the temperature of tissue to a therapeutic level. For ablation purposes, coagulative necrosis is usually achieved by heating the tissue above ~45 °C. The amount of absorption depends upon the nature of the tissue, its degree of vascularization, and the frequency and intensity of the applied ultrasound. Tissues with high protein content absorb ultrasound more readily than those with high fat content. Thus, tissue with high water content and low protein such as blood and fat content absorb little of the US energy while those with lower water content and higher protein content such as ligament and tendon will absorb US more efficiently. Although cartilage and bone have the highest protein content, their densities cause problems with wave reflection and a significant proportion of US energy striking their surfaces is more likely to be reflected. The best absorbing tissues in terms of clinical practice are those with high collagen content: ligament, tendon, fascia, joint capsule, and scar tissue [ter Haar, 2007; Watson, 2000]. Another key point is that the higher the US frequency, the greater the absorption rate. Raising the temperature above normal thermal levels by a few degrees may have a number of beneficial physiological effects [ter Haar, 2007]. Most common frequencies used are in the range from 0.7 to 3.3 MHz [ter Haar, 2007]. Maximum energy absorption depth in soft tissue is 2 to 5 cm [Watson, 2008].

Non-thermal effect: Therapeutic ultrasound produces a combination of non-thermal effects that are difficult to isolate from the thermal [Johns, 2002]. Non-thermal effects have been divided by ter Haar into cavitations and other mechanical effects such as acoustic streaming and microstreaming [ter Haar, 2007]. Cavitations refer to the behavior of bubbles within an acoustic field. They are defined as the physical forces of the sound waves on micro-environmental gases within a fluid. As the sound waves propagate through the medium, the characteristic compression and rarefaction causes microscopic gas bubbles in the gas to contract and expand. The thought is that the rapid changes in pressure in and around the cell may alter the function of the cell. Acoustic streaming has been described by ter Haar as “localized liquid flow in the fluid around the vibrating bubbles” [ter Haar, 2010]. This has been further defined as the physical forces of the sound waves that provide a driving force capable of displacing ions and small molecules. At the cellular level, organelles and molecules of different molecular weight exist. Many of these are free floating and may be driven to move around more stationary structures [Johns, 2002]. Microstreaming is set up in the fluids around acoustically driven bubbles. This purports to lead to shear stresses on cell membranes in the vicinity, which can create transient pores through which ions and molecules may be transported [ter Haar, 2007]. This increased permeability of both individual cell membranes and the endothelium is thought to enhance therapeutic uptake, and can locally increase the activity of drugs by enhancing their transport across membrane [Stride & Coussios, 2010]. Recently, a continuum mechanics hypothesis has been put forward, stating that ultrasound can noninvasively modulate neuronal activity through a combination of pressure/fluid/membrane actions involving stable cavitation and acoustic streaming (microjet formation, eddying and turbulence) in addition to acoustic radiation force, shear stress, Bernoulli effects and other fluid-mechanical consequences, which stem from small acoustic impedance mismatches (boundary conditions) between lipid bilayers, surrounding intracellular/extracellular fluids, and interleaved cerebrovasculature [Tyler et al. 2010].

**Acoustic Spectrum**

Audible sound is what humans hear in the approximate frequency range between 20 Hz and 20 kHz. The ultrasound frequency range starts at a frequency of about 20 kHz. Most medical
equipment operates in the ultrasonic frequency range between 1 to 15 MHz [O’Brien, 2007]. Therapeutic ultrasound has a frequency range of 0.75-3 MHz, with most machines set at a frequency of 1 or 3 MHz. Low-frequency ultrasound waves have a greater depth of penetration but are less focused (although novel phase-corrected transducer arrays are capable to enhance the focus [Clement & Hynynen, 2002; Hynynen et al., 2004]). For physical therapy, ultrasound at a frequency of 1 MHz is absorbed by tissues at a depth of 3-5 cm and is recommended for deeper injuries and in patients with more subcutaneous fat. A frequency of 3 MHz is recommended for more superficial lesions at depths of 1-2 cm [Speed, 2001].

**Thermal Dose**

Healthy cellular activity depends upon chemical reactions occurring at the proper location at the proper rate. The rate of chemical reactions and thus of enzymatic activity are temperature dependent. An immediate consequence of a temperature increase is an increase in biochemical reaction rates. When the temperature becomes sufficiently high (i.e., above ~45 °C), enzymes denature [O’Brien, 2007]. With this in mind, therapeutic applications require that the exposed target tissue undergoes reversible or irreversible change, depending on the goal of the treatment [ter Haar, 2007]. Application in a clinical environment is determined by a combination of the selectable machine parameters: frequency, power density, duty cycle and treatment time. Ultrasound dosages are also varied by alteration of wave amplitude and intensity (W/cm²), which can be pulsed or continuous. Continuous ultrasound has a greater thermal effect but either form at low intensity will produce non-thermal effects [Speed, 2001]. Studies utilizing continuous and pulsed frequencies at 1MHz and 3 MHz confirm that ultrasound results are time and dose dependent. The 3 MHz frequencies increased tissue temperature at a faster rate than the 1MHz frequency [Johns, 2002]. For example, a 2004 in vivo study concluded that pulsed ultrasound (3MHz, 1.0 W/cm², 50% duty cycle, for 10 minutes) produces similar intramuscular temperature increases as continuous ultrasound (3MHz, 0.5 W/cm², for 10 minutes) at a 2-cm depth in human gastrocnemius muscles [Gallo et al., 2004].

**Coupling Media**

Sound waves are transmitted through a round-headed wand that is applied to the skin; however, the waves may be reflected at the metal/air interface found at the treatment head. It is therefore necessary to provide a medium through which the ultrasound can pass freely to reach the tissue without absorbing or changing the direction of the waves. This medium should be sufficiently fluid to fill all available spaces, relatively viscous so that it stays in place, have impedance appropriate to the media it connects, and should allow transmission of ultrasound energy with minimal absorption, attenuation or disturbance [Watson, 2008]. For therapeutic applications, a number of methods are used to couple the sound into the tissue. Where the acoustic window is relatively flat, and the transducer’s emitting surface is flat, aqueous gel may be used between the source’s front face and the skin. For irregular tissue surfaces and/or irregular transducers, water may provide a better coupling medium [ter Haar, 2007]. In 2004, Casarotto et al compared 4 coupling media: gel, mineral oil, white petrolatum, and degassed water for density and temperature variation. The results showed that the water and gel presented the highest transmission coefficient, the lowest reflection, and an attenuation coefficient and acoustic impedance close to that of the skin. However, when using direct contact and thin layers of coupling agents, any product may be used, because the effect of the attenuation coefficient does not play a significant role when layers are very thin [Casarotto, 2004].

**Calibration and Safety**

The calibration standard for ultrasonic power output is monitored by the FDA code of federal regulation title 21, part 1050.10 which states that temporal-average ultrasonic power shall not exceed ± 20% for all emissions greater that 10 % of the maximum emission [Food and Drug Administration, Code of
Federal Regulations, 2009]. The IEC standard for physiological equipment includes two limits for the purpose of patient protection. The first limits the temperature of the frontal face of the transducer to no more than 41°C when operated under water with initial temperature 25 °C. A test involving three 3-minute cycles is specified. The second limit applies to ultrasound intensity. The effective intensity shall not exceed 3 W/ cm². Extended exposure of tissue to this intensity causes temperature increases, which can result in tissue damage, particularly at the surface of bone. Protection of non-target tissue regions is achieved from appropriate placement of the beam [Duck, 2007].

4 Applications of Therapeutic Ultrasound

Ultrasound is relatively well established as a treatment modality for a number of conditions ranging from soft tissue lesions to kidney stones and cancer. Understanding the use and limitations of ultrasound in the treatment of these conditions will help in the development of ultrasound treatment targeted to select brain regions.

Lithotripsy

In extracorporeal shock wave lithotripsy (ESWL), up to 120 high-energy ultrasonic pulses are delivered

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1880</td>
<td>Piezoelectric effect (Curie)</td>
</tr>
<tr>
<td>1907</td>
<td>Electronic Vacuum tube (Deforest)</td>
</tr>
<tr>
<td>1918</td>
<td>Sonar (Langevin)</td>
</tr>
<tr>
<td>1927</td>
<td>Effects on biological tissues (Looms, Wood)</td>
</tr>
<tr>
<td>1942</td>
<td>HIFU effects (Lynn, Putnam)</td>
</tr>
<tr>
<td>1950-1969</td>
<td>Molecular studies on HIFU effect (Fry)</td>
</tr>
<tr>
<td>1951-1960</td>
<td>Radiofrequency generator and electrodr development (Cosman, in light of FUS developments)</td>
</tr>
<tr>
<td>1951-1967</td>
<td>Radiosurgery and gamma knife development (Laksell after ultrasound investigation)</td>
</tr>
<tr>
<td>1960-1980</td>
<td>Clinical studies on HIFU surgery with open cranium</td>
</tr>
<tr>
<td>1980s - present</td>
<td>MRI-technology</td>
</tr>
<tr>
<td>early 1990s</td>
<td>Ultrasound phases arrays (Hynynen)</td>
</tr>
<tr>
<td>mid-1990s</td>
<td>MRI thermometry (Jolesz)</td>
</tr>
<tr>
<td>2001</td>
<td>First integrated MRgFUS machine (Insightec)</td>
</tr>
<tr>
<td>2006</td>
<td>Report on MRgFUS for treatment of GBM after craniotomy (Ram)</td>
</tr>
<tr>
<td>Present</td>
<td>Trial on MRgFUS with cranium intact (Jolesz)</td>
</tr>
</tbody>
</table>

Figure 1. Timeline of the Development of High-intensity Focused ultrasound technology

through acoustic focusing lenses to break calculi such as kidney stones and gallstones into fragments small enough to be passed from the body without undue difficulty [Terhorst et al., 1972; Pfister et al. 1988; Madersbacher & Marberger, 2003]. The successive shock wave pressure pulses result in direct shearing forces, as well as cavitation bubbles surrounding the stone, which fragment the stones into smaller pieces. The shock waves themselves, as well as cavitation bubbles formed by the agitation of the urine medium, can lead to capillary damage, renal parenchymal or subcapsular hemorrhage [ter Haar, 2007]. This can lead to long-term consequences such as renal failure and hypertension. Overall complication rates of ESWL range from 5–20% [Auge & Preminger 2002; Madaan & Joyce, 2007].

**Physical therapy**

Ultrasound (pulsed and continuous, 0.1-3 W/cm², 0.5-3 MHz) is commonly used in the orthopedic setting to decrease soft tissue inflammation, increase tissue extensibility, enhance scar tissue remodeling, increase soft tissue healing, decrease (muscle and joint) pain, decrease soft tissue swelling and to deliver medication for soft tissue inflammation [Wilkin, et al., 2004; Wong, et al., 2007]. However, recent clinical trials are in disagreement about the benefit of ultrasound therapy in musculoskeletal conditions. Some reports are in clear support of ultrasound therapy [Binder et al.,1985; Piravej & Boonhong, 2004; Bakhtiary & Rashidy-Pour, 2004; Majlesi & Unalan, 2004; Baysal et al., 2006; Srbely & Dickey 2007; Dincer et al., 2009], whilst others seem to suggest that ultrasound therapy does not actually confer any advantage to established treatment regimes [van der Windt et al., 2002; Gam et al., 1998 ; van der Windt et al., 1999; D’Vaz et al., 2006].

**Drug delivery**

Through its mechanical forces and heat generation, ultrasound has been shown to enhance the permeability of cells to therapeutic drugs. This property of ultrasound has been successfully used in so called acoustic targeted drug delivery (ATDD) [Lewis et al. 2008]. Furthermore, recent studies show that HIFU can selectively disrupt the blood-brain barrier (BBB) [Mesiwala et al., 2002; Rajadhyaksha et al., 2011; Tung et al., 2011; Madsen & Hirschberg, 2010; Hayner & Hynynen 2001; Hynynen et al. 2006, Hynnen 2008; Choi et al. 2007; Liu et al. 2009; Xia et al. 2009]. It has been argued that selective manipulation of the BBB may hold great promise for the treatment of brain disorders as it (1) enables the administration of existing drugs that are currently ineffective in the brain, (2) allows for the development of new drugs without the restraints imposed by the BBB and (3) enhances the prospect of targeted drug delivery to desired brain regions, reducing systemic toxicity [McDannold et al. 2011]. However, there is mounting evidence that HIFU induced BBB disruption is associated with a number of adverse effects, including leakage of plasma proteins into the brain (albumin is toxic to astrocytes [Nadal et al., 1995], and astrogliosis is induced when brain comes in contact with blood), vascular pathology [Treat et al., 2007; Yang et al., 2007; Lossinsky et al., 1995] and chronic neuropathologic changes in the brain [Salahuddin et al., 1988].

**High Intensity Focused Ultrasound**

Using lower frequencies (250–2000 kHz) than medical diagnostic ultrasound but significantly higher time-averaged intensities (typically more than 1000 W/cm²), High Intensity Focused Ultrasound (HIFU), also called Focused Ultrasound Surgery (FUS), has been successfully used to ablate tumors or other tissue through coagulative necrosis. Conditions currently treated with HIFU include uterine fibroids [Ren et al., 2007], prostate cancer [Lindner et al. 2010] and other cancers [Ng et al., 2011; Orsi et al., 2010]. This kind of treatment is often guided by Magnetic Resonance Imaging (MRI)-based thermometry, which provides constant feedback of the ultrasound beam location and focusing in a closed-loop system. The combination of the two is referred to as Magnetic Resonance-guided Focused Ultrasound Surgery (MRgFUS) [Mikami et al., 2008; Jolesz, 2009].
Focused ultrasound was first used in brain surgery as early as in the 1950s, when Fry and colleagues demonstrated the use of high intensity focused ultrasound to produce focal destructive lesions [Wall et al., 1951; Fry et al., 1954, 1955]. Since then, the use of FUS in brain surgery was limited by the inability to target and monitor treatments as well as the fact that it required invasive craniotomies thereby offering little advantage over straightforward surgical interventions. Nevertheless, following preliminary experience with MRI-guided thermoablation of cerebral neoplasms [Anzai et al., 1995], Ram and colleagues published the first clinical report about MRI-guided focused ultrasound as a potential modality to destroy tumor tissue by thermocoagulation through a small craniotomy [Ram et al., 2006]. In the meantime, technological advances have made it possible to transmit and focus ultrasound through the intact human skull using an array of phase-corrected ultrasonic transducers placed on the cranium [Clement & Hynynen, 2002; Hynynen et al., 2004]. Using this technology, McDannold and colleagues demonstrated the feasibility of transcranial MRI-guided HIFU (tcMRgHIFU) in three patients and visualized the heating with MRI temperature imaging in a noninvasive fashion [McDannold et al., 2010]. Using a higher power instrument, it has recently been shown that tcMRgHIFU is effective in the non-invasive thermal ablation of thalamic nuclei in neuropathic pain patients [Martin et al., 2009]. In this phase 1 clinical trial the central lateral thalamic nuclei in nine patients with chronic pain was successfully heated to more than 53 °C to achieve thermal ablation. All patients tolerated the fully awake noninvasive procedure in an outpatient setting without any side effects or neurologic deficits. In all treated patients, pain relief to some extent was immediate lasting until publication, although prior studies of thalamotomy have suggested that this relief might be only transient. These results and continuously improving methodology promise to rekindle interest in the use of ultrasound in neurosurgery [Tyler, Tufail & Pati, 2010; Jagannathan et al. 2009].

Low-Intensity Low-Frequency Ultrasound

In addition to the destructive lesions produced by FUS, Fry and colleagues also showed that lower intensity ultrasound can produce reversible suppression of activity in the central nervous system [Fry et al. 1958]. This was followed later by studies demonstrating the use of low intensity ultrasound to stimulate nervous activity through the skin. Gavrilov and colleagues showed that short, focused ultrasound stimuli could induce a variety of somatic sensations, including tactile, thermal (warm and cold) and pain, without causing damage to the treated tissue [Gavrilov et al. 1977; Gavrilov et al., 1980]. These results encouraged the use of low intensity ultrasound in the treatment of joint pain in humans [Wright et al, 1993] and have since been used to treat a various musculoskeletal conditions (see above, 3.2.2). Low intensity ultrasound can also induce auditory responses when applied to the cranium in animals [Foster & Wiederhold, 1978] and humans [Gavrilov et al. 1977; Gavrilov et al., 1980; Magee & Davies, 1993]. These sensory effects are thought to be mainly mediated by non-thermal, mechanical effects on neurons and resulting changes in membrane potential. Tyler and colleagues suggest that the modulation of ionic conductance produced by adiabatic processes as ultrasound propagates rapidly and transiently through neuronal membranes may alter the activity of individual neurons due to the elastic nature of lipid bilayers and the spring-like mechanics of many transmembrane protein channels [Tyler et al., 2008]. In partial support of this hypothesis, low-power ultrasound has been shown to influence the membrane conductance of frog skin epidermis [Dinno et al., 1989]. In addition, ultrasound exposure can induce a reversible increase in the internal Ca2+ concentration of fibroblasts [Mortimer & Dyson, 1988]. In rat thymocytes, stimulation with ultrasound can modulate K+ influx and efflux [Chapman, MacNally & Tucker, 1980]. Several investigations have demonstrated that ultrasound modulates neuronal activity by enhancing and/
or suppressing the amplitudes and/or conduction velocities of evoked nerve potentials [Mihran, Barnes & Wachtel, 1990; Rinaldi et al., 1991; Bachtold et al., 1998; Tsui, Wang & Huang, 2005]. By changing intensity and duration of focused ultrasound nerve fibers of different diameter and myelination can be differentially affected [Young & Henneman, 1961]. Given that many voltage-gated ion channels, as well as neurotransmitter receptors possess mechanosensitive properties that render their gating kinetics sensitive to transient changes in lipid bilayer tension it is conceivable that ultrasound directly modulates ion channel activity [Sukharev & Corey, 2004; Morris & Juranka, 2007]. A recent study monitoring calcium signals in cultured hippocampal slices showed that low-intensity low-frequency ultrasound (LILFU) indeed activates voltage-gated sodium channels and stimulates neuronal calcium channels, whilst having no adverse effect on the fine structure of neuronal membranes even following chronic LILFU stimulation [Tyler et al., 2008].

5 Conclusions

Recent advances in the development of transcranial focusing methods have rekindled interest in therapeutic ultrasound for the treatment of brain disorders. Currently, a number of clinical trials are under way investigating the use of this improved technique in the treatment of brain cancer and pain. These studies employ transcranial ultrasound to make discrete lesions.

However, using these novel transducer arrays and phase-correction algorithms, it should also be possible to make non-invasive, reversible manipulations to brain structures using low-intensity and/or low-frequency ultrasound. Recent results from Tyler and colleagues (2008) encourage this idea. The group demonstrated that low ultrasound frequencies (0.44 – 0.67 MHz) affect ion channel opening and synaptic release in cultured neurons whilst having no other adverse effects. It should be noted that the frequencies used match the optimal gain between transcranial transmission and brain absorption (~0.60 – 0.70 MHz) [Hayner & Hynynen, 2001; White, Clement & Hynynen, 2006].

In addition to allowing non-invasive, reversible manipulations of brain function, the mobility and comparatively low cost of transcranial therapeutic ultrasound equipment make it a potentially highly popular treatment modality in future.

6 Future Work

Recent work by Tyler and colleagues (2008) suggests that a non-invasive, reversible brain dysfunction treatment modality employing therapeutic ultrasound is feasible, and there are currently a number of clinical trials underway investigating the use of this methodology. Additional development and study of innovative transducer arrays and phase-correction algorithms will likely yield beneficial results. The goal of proponents of transcranial therapeutic ultrasound is to further refine this technology and methodology to a point that will make it a reliable, fully non-invasive and reversible treatment for brain dysfunction.

Acknowledgement

The authors want to acknowledge Dr. Michael Kohl as a co-author of this work, for his outstanding contribution of original research and literature review. The authors wish to also acknowledge Dorian Bakogiannis and Mark Atallah for editing and communication support in the preparation of this paper.

References


Submission Invitation

The Brain Sciences Journal supports cross-disciplinary articles relating to cognition and neuropathology. The BSJ welcomes research from a wide variety of backgrounds including neuroscience, psychology, mathematics, biology, computation, and philosophy. The BSJ publishes papers on the study of mind and or brain and nervous system in all applications ranging from theoretical to clinical.

Topics include, but are not limited to:

- Alzheimer’s Disease
- Artificial Intelligence
- Bioinformatics
- Brain Patterning
- Brain/Computer Interface
- Brain/Mind Duality
- Cognition
- Cognitive Aging
- Cognitive Behavior
- Cognitive Development
- Cognitive Ergonomics
- Cognitive Modeling
- Cognitive Neuroscience
- Cognitive Psychology
- Cognitive Therapy
- Collective Intelligence
- Computational Linguistics
- Decision Making and Reasoning
- Ethics and Policy Issues in Neuroscience
- Human Learning
- Human/Computer Interaction
- Imagination
- Intelligence
- Language
- Language Acquisition
- Learning
- Machine Learning
- Mechanisms of Attention
- Memory
- Natural Language Processing
- Neurolinguistics
- Neurological Disorder
- Neurophilosophy
- Perception
- Philosophy of Mind
- Psychology
- Robotics
- Social Cognition
- Speech Processing Synthetic Intelligence
- Theory of Computation
- Unconscious Mind
- Working Memory

Submission Guidelines:

1. **Primary Research Reports:** Full articles and brief communications both of which are restricted to 10,000 words overall, with exception. Articles include but are not limited to, toolboxes, which provide a critical overview of a method or technology that is becoming widely used. Reviews, which provide an overview of a single theme or topic for the broad brain sciences audience. Disease focused articles, which provide brief overviews of a neural disease or syndrome and describe how that disorder might be linked to basic neural mechanisms.

2. **Analyses:** Analyses articles offer a comprehensive comparative examination of technologies, methods or reagents of key importance for a field of research, leading to important practical conclusions about their performance. Analyses articles may also report a new analysis of existing large datasets that lead to a novel, exciting or arresting conclusion.

3. **Visions:** The Brain Sciences Journal will publish papers that present new ideas, theories and paradigms for the field of brain sciences. Vision papers are usually no longer than 15,000 words, including figure legends and have up to 100 references.

For more information on how to prepare your manuscript please visit: [www.brainsciencesjournal.org/submissions](http://www.brainsciencesjournal.org/submissions)

Submission:

Please send your manuscript as an attachment to: rebecca.fahlstrom@brainsciences.org

* For more information please visit [www.brainsciencesjournal.org](http://www.brainsciencesjournal.org)
* If you have any questions please contact rebecca.fahlstrom@brainsciences.org