A Blueprint for the Hard Problem of Consciousness

Potential States, Strange Loops and The Realization of Information from Cells to Human Consciousness

CONTENTS

FOREWORD i

WHY WE NEED A BLUEPRINT FOR CONSCIOUSNESS i
THE PROBLEM OF CONSCIOUS VERSUS UNCONSCIOUS PROCESSES ii
THE PROBLEM OF THE UNANALYZABILITY OF CONSCIOUS CONTENTS iii
THE PROBLEM OF COMPLETENESS iv
THE PROBLEM OF CONTEXTUAL SENSITIVITY iv
THE PROBLEM OF THE SPEED OF PROCESSING v
THE NEED FOR A BLUEPRINT OF CONSCIOUSNESS vi
NOTE vii

PREFACE viii

CONSENT FOR PUBLICATION xii
CONFLICT OF INTEREST xii
ACKNOWLEDGEMENTS xii
NOTE xiii

DEDICATION xiv

SECTION 1. SETTING UP THE PROBLEM

CHAPTER 1 INTRODUCTION 1

WHAT IS THE HARD PROBLEM OF CONSCIOUSNESS 1
CONSCIOUSNESS AS A FORM OF INFORMATION 2
THE DIFFERENCE BETWEEN NATURAL HISTORY AND EVOLUTIONARY THEORY 2
NOTES 3

CHAPTER 2 DOES IT MAKE SENSE TO TRACE CONSCIOUSNESS TO SINGLE-CELLED ORGANISMS? 4

QUALIA AS INFORMATION 6
NOTES 8

CHAPTER 3 CONSCIOUSNESS: COMPUTABLE, NONCOMPUTABLE OR BOTH? 10

IN FAVOR OF NONCOMPUTABILITY 10
IN FAVOR OF COMPUTABILITY 11
A HYBRID SYSTEM: EMBODIED ABSTRACTIONS 13
NOTES 16

CHAPTER 4 A BRIEF CRITICISM OF THEORIES 19

THE CIRCULAR REASONING OF THE INTEGRATED INFORMATION THEORY 19
THE PROBLEM LIES IN THE IIT AXIOMS 21
QUANTUM INFORMATION, ANTHROPOMORPHIZED IN THE ORCH OR THEORY 21
THE ABSENT SUBJECT 23
NOTES 23

SECTION 2. BLUEPRINT

CHAPTER 5 EXPECTATION WITHOUT A SPECTATOR 25

INFORMATION IS MADE POSSIBLE BY EXCHANGES BETWEEN THE ORGANISM AND ITS ENVIRONMENT 25
A CAR ENGINE EXPRESSES ITS DESIGNERS' EXPECTATIONS 26
LIVING BEINGS EXPRESS EMERGING PROCESSES EXPECTATIONS 26
NOTES 28

CHAPTER 6 NEGATIVE ENTROPY AND INFERENCE IN LIVING ORGANISMS 30

CELLS TRANSFORM SUBCELLULAR ORGANIZATION IN BIOLOGICALLY MEANINGFUL INFORMATION WITH COMMUNICATIONS CHANNELS AND REFERENCE ABOUT SOMETHING 30
CELLULAR PROCESSES AND BRAIN COMPUTATIONS SHARE AN IMPLICIT PROBABILISTIC MODEL OF THE ENVIRONMENT 31
NOTES 33

CHAPTER 7 EXPECTATIONS IN A STRANGE LOOP 36

THE INTRODUCTION OF 'EXPECTATION.INFORMATION' EXPECTATION.INFORMATION 36
A GENERATIVE MODEL OF CONSCIOUSNESS IS BASED ON ACTIONS 38
NOTES 40

CHAPTER 8 IN SEARCH OF MEANING 41

SHANNON'S INFORMATION CAN BE REALIZED AS INFORMATION ABOUT SOMETHING 41
NOTES 44

CHAPTER 9 SUPPLANTING THE INTEGRATED INFORMATION THEORY 45

HUMAN SUBJECTIVITY IS ABSTRACT AND CANNOT BE MEASURED DIRECTLY 45
THE IIT IS A CRYPTIC FORM OF INTELLIGENT DESIGN 47
NOTES 48

SECTION 3. IN SEARCH OF QUALIA IN LOCAL NETWORKS

CHAPTER 10 AWARENESS OF LOCATION 49

A BRIEF DESCRIPTION OF THE PHYSIOLOGY 50
CAN QUALIA BE PRESSED DOWN THE RABBIT HOLE OF LOCAL NETWORKS? ... 51
REALIZATION OF INFORMATION ACROSS OMNIDIRECTIONAL CELL

ASSEMBLIES CAN CONVEY ABSTRACT MEANING 52
THETA OSCILLATIONS CROSS-COUPLING CAN GENERATE HIGHER-ORDER CELL ASSEMBLIES 53
ACTIONS GENERATE CONCRETE AND ABSTRACT MAPS 55
CONCLUSIONS 56
NOTES 56

CHAPTER 11 AWARENESS OF TIME 60

A BRIEF REVIEW OF THE PHYSIOLOGY 61
AN INTRODUCTION TO INTERVAL TIMING 61
LOCAL NETWORKS CAN ENCODE LOCAL TIME 62
SHORT-TERM PLASTICITY CAN EXPLAIN LOCAL EFFECTS IN INTERVAL TIMING 62
OSCILLATORY PHASE ADAPTATION FRAMES BRAIN-CENTRIC TIME EXPERIENCES 63
BRAIN-CENTRIC EXPECTATIONS PREDICT CONSCIOUS PER- CEPTION 64
CONCLUSIONS 64
NOTES 65

CHAPTER 12 INTEROCEPTION 69

A BRIEF REVIEW OF THE ANATOMY AND PHYSIOLOGY OF THE INSULA 70
THE EMBODIMENT OF ONE'S SENSE OF SELF 71
THE REPRESENTED BODY AND THE REPRESENTED SELF 71
A HIERARCHY OF ABSTRACTIONS IN THE INSULA 72
CONCLUSIONS 73
NOTES 73

CHAPTER 13 PREDICTIVE CODING 76

PREDICTIVE CODING SUPPORTS A SUBJECT-CENTRIC CONSCIOUSNESS 77
PREDICTIVE CODING IS CONSISTENT WITH BAYESIAN INFERENCE OF HIDDEN CAUSES IN THE ENVIRONMENT 78
HOMO HEURISTICUS 81
CONSCIOUSNESS, FAST AND LOW 82
CONCLUSIONS 84
NOTES 85

CHAPTER 14 CORTICAL SIGNALS OF CONSCIOUSNESS 89

A BRIEF REVIEW OF THE PHYSIOLOGY 89
SIGNALS OF CONSCIOUSNESS IN CORTICAL LAYER 5 90
CONCLUSIONS 91
NOTES 92

CHAPTER 15 OLFACTORY AWARENESS 94

A BRIEF REVIEW OF THE PHYSIOLOGY 95
FROM ODOR RECEPTORS TO OSCILLATIONS IN THE OLFACTORY BULB 95
EARLY AND LATE OLFACTORY GAMMA OSCILLATIONS ARE CONSISTENT WITH EXPECTATION.INFORMATION 96
INFORMATION FLOW ACROSS THE PIRIFORM CORTEX AND THE EMERGENCE OF "ODOR-OBJECTS" 97
THE MATCHING OF OLFACTORY AND INTEROCEPTIVE HIERARCHY OF ABSTRACTIONS 100
THE SYNTHESIS OF SMELL IS CONSISTENT WITH QUALIA AS INFORMATION ... 101
CONCLUSIONS 104
NOTES 106

SECTION 4. THE GLOBALIZATION OF CONSCIOUSNESS

CHAPTER 16 ATTENTION, NEURONAL OSCILLATIONS AND PREDICTIVE CODING 113

ATTENTIONAL MODELS 114
INTERNAL MODELING 115
FIRST ORDER AND HIGHER ORDER REPRESENTATIONISM 115
COMPUTATIONAL MODELS AND PREDICTIVE CODING 116
BETWEEN THE EASY AND THE HARD PROBLEMS 117
A BRIEF SUMMARY OF NEURONAL OSCILLATIONS 117
INTEGRATION OF LOCAL AND GLOBAL ACTION 120
LOCAL AND GLOBAL INTERACTIONS IN VISUAL AWARENESS 120
REVISION OF CORTICAL SIGNALS OF CONSCIOUSNESS WITH INTRACRANIAL ELECTRODES MEASUREMENTS: A HYBRID LOCAL AND GLOBAL MODEL 121
"BEING THE SUBJECT" IS THE REALIZATION OF INFERENCE INVERSELY MAPPED OUT AS HIDDEN CAUSES OF GLOBALLY INTEGRATED ACTIONS 123
THE LOCALIZATIONISM ERROR OF MY WAYS 124
THE LIMITATIONS OF THE LABELED-LINE CONCEPT 124
CONCLUSIONS 125
NOTES 126

CHAPTER 17 THE GLOBAL NEURONAL WORKSPACE THEORY AND THE ADAPTIVE RESONANCE THEORY 132

THE GLOBAL NEURONAL WORKSPACE THEORY 133
A BRIEF REVIEW OF THE THALAMIC PHYSIOLOGY 134
PARALLELS BETWEEN SILICON AND BRAIN COMPUTATION 135
GLOBAL AND LOCAL CONNECTIVITY AS A FORM OF INFORMATION 136
SUBJECT-CENTRIC SYNCHRONIZATION 137
ABSTRACT STATES OF AWARENESS DECOUPLED FROM QUALIA 137
THE ADAPTIVE RESONANCE THEORY 138
CONCLUSIONS 139
NOTES 140

CHAPTER 18 NETWORKS OF EXTERNAL AND INTERNAL AWARENESS AND THE NEUROLOGY OF CONSCIOUSNESS 145

THE DEFAULT MODE NETWORK 146
RECONCILING THE GLOBAL NEURONAL WORKSPACE AND THE DEFAULT MODE NETWORK 147
ANESTHESIA STUDIES 148
FEEDBACK CONNECTIVITY IN ANESTHESIA 151
EFFECTS OF ANESTHETICS ON PYRAMIDAL NEURONS APICAL DENDRITES 152
QUANTUM OPERATIONS IN DENDRITES 155
CONSCIOUSNESS WITHOUT A CORTEX 156
THE SPRAGUE EFFECT 157
THE BRAIN STEM CONVERGENCE SUPPORTS A PHENOMENAL SUBJECT-CENTRIC MOTOR ANALOGIC SIMULATION 158
INTEGRATION FOR ACTION 159
THE UPPER BRAINSTEM MAY SUPPORT PREDICTIVE CODING 160
EVIDENCE OF CONSCIOUSNESS IN ANENCEPHALIC CHILDREN CAN MAKE THE GNW A FALSIFIABLE THEORY FOR THE HARD PROBLEM OF CONSCIOUSNESS 160
CONCLUSIONS 161
NOTES 162

SECTION 5. THE ROAD AHEAD

CHAPTER 19 A PROBLEM OF RECORDS 169

WHY IS A QUESTION ABOUT RECORDS IMPORTANT? 169
THE PROBLEM OF RECORDS IS PHYSICAL, NOT METAPHYSICAL 172
THE PROBLEM OF RECORDS IS NOT ALGORITHMICALLY COMPRESSIBLE 172
COMPETING EXPLANATIONS OF CONSCIOUSNESS IMPLY DIFFERENT VERSIONS OF THE UNIVERSE 174
THE ABSENT SUBJECT 175
NOTES 177

CHAPTER 20 FINAL CONCLUSIONS 180

A BRIEF SUMMARY 181
CONCLUSIONS 187
SOME PERSONAL NOTES 189
NOTES 192

REFERENCES 193

PREFACE

The origin of subjectivity constitutes one of the most challenging scientific questions in any field of human research. The fundamental mechanism that allows physical events to transcend into subjective experiences continues to elude scientists and philosophers, who have aptly named it the Hard Problem of Consciousness.

This blueprint is an abstract model that I envisioned as an explanation for the origin of subjectivity. In many aspects it is still a crude idea demanding testing and validation. It is best seen as a compass, or a map, indicating a direction towards the solution of the problem.

Here, I propose how subjectivity could have arisen out of nothing or, at least, nothing of its own kind. Human beings developed through the course of a conservative natural history. Thus, a connection must exist between the processes that sustain the life of primitive single cells, and the emergence of human consciousness. Moreover, such an association could only have taken place if these ancient biological processes, at some point, started to convey information. In other words, single-celled organisms and their communication channels with the environment generated information, as defined by Shannon's classical information theory.

Current theories of consciousness suffer from an intelligent design bias. Their proponents often search for answers in patterns of information integration, in quantum models, and in other natural phenomena. The subject, a necessary vessel for the existence of subjectivity, becomes an afterthought, and consciousness penetrates every aspect of the physical world, in a cryptic panpsychism.

In contrast, I took the subject's perspective, and looked for a solution to the Hard Problem in the boundaries that separate living subjects from non-living processes. Under this light, consciousness emerges as a form of realization of information, connecting simple life forms to their environment. Consciousness in single-celled organisms, in neural networks, or in the intact human brain shares the same mystery: the nothingness or emptiness of its existence. How could something be there and not be there at the same time?

The realization of information must occur as a self-referential abstraction, otherwise consciousness would not differ from other physical properties of inanimate objects, such as

heat.

I use the word “abstraction” to portray a confluence of concepts. In computer science, an abstraction enables programmers to handle the complexity of systems by subsuming details in lower levels. However, there is no icon in a computer screen depicting qualia. The brain extracts and processes regularities from its environment through neuronal assemblies within progressively higher orders. These assemblies have a relational origin, describe phenomenal objects, and can be detached from their conditions of origin. I call this extraction by inference an “abstraction”.

Finally, I consider the realization of self-referent information across complexity levels an abstraction. When high-level realization of information occurs by means of a strange loop, this abstraction of self-referent origin can be made available to the organism as a form of subjectivity. Thus, an abstraction is not physically there, despite having a mathematical description. The physical components and interactions that give rise to abstractions can be measured, including the activity of neuron assemblies assessed by electrophysiological experiments or functional brain neuroimaging. Nevertheless, the realization of information cannot be measured directly. It is a form of nothingness.

Consider the following analogy. Capital emerges out of nothing of its order. The concept is well accepted in economic science. Capital is something abstract, not physically present. It is a consequence of the economic dialectic of companies and the market. It exerts an organizing effect on the economic activities of the same companies. Capital can be represented as a variable in economic theory. Capital has a physical and an informational basis, such as in real estate, machine tools, patents, cash, and the know-how of the working force. Capital can be realized as money. Note that paper money also represents a concept: a measure of value, a promise to pay the bearer. In the context of this book, to realize is to make it available at a different level and organization.

I use the word realize to express the additional meaning of “knowing[2]”. To realize is the act of knowing. The present work shows what sort of mechanisms could render an abstract emergent order, such as capital, aware of itself.

Here, I propose a new version of Hofstadter’s strange loop that allows for the realization of information between actual and potential states. I call such states the “information” and “expectation” of processes that connect an organism and its environment.

The *information.expectation* relationship underlies an inferential system. This implicit probabilistic model of the environment is encoded in the structure and processes of single celled organisms and shares the same inferential informational paradigm of abstract information realized across the brain computational hierarchy as the strange loop:

‘information.expectation’information.expectation

I studied how local brain networks relate to experiences of awareness. Specifically, I drew ideas from work on spatial localization done on the hippocampus, but also from studies on the awareness of time, interoception – the awareness of our internal body, and the olfactory system.

At this level of organization, the brain seems to follow the rules of something called Predictive Coding. The brain does not build perceptions from the bottom up; it creates hypotheses or beliefs from the top down. In other words, Predictive Coding supports the idea of a subject-centric consciousness, and allows for the following reformulation of the strange loop:

'expectation.information'expectation.information

which in the feedback and feedforward language of Predictive Coding of corticocortical and thalamocortical networks can be rewritten as

'top-down.bottom-up'top-down.bottom-up

or

'prediction.predictive-error'prediction.predictive.error

Since the beginning, I believed in a noncomputational origin for consciousness, and built up my thought experiment around this hypothesis. At its roots brain activity is analogic, not digital. Surprisingly, however, the problem of locally encoded consciousness can be approached from a computational perspective.

Computation in both silicon chips and the brain follows a hierarchy that inherently supports complex systems. Thus, consciousness conceptually emerges as the realization of information, as a hierarchical abstraction through a new interpretation of strange loops across abstraction levels.

This solution explains how local networks, while devoid of consciousness, encode the information necessary for the emergence of consciousness. Furthermore, the idea of consciousness as a computationally derived abstraction fits with the concept of something immaterial springing out of a physical substrate.

This model predicts that the realization of information as some form of awareness requires sustained neuronal rhythms that depict expectation-based Predictive Coding. The awareness of simple phenomenal experiences becomes possible when these oscillatory rhythms integrate sensory information in a pre-motor framework for sequential global actions. Being the subject is the realization of inference inversely mapped out as hidden causes of globally integrated actions. Hofstadter's strange loop applies here.

This action-oriented "generational" model bears computational and noncomputational aspects. It predicts the generation of hypotheses by the brain across multiple scales of cross-frequency oscillatory coupling. A model based on the generation of hypotheses constitutes a simulation in itself. Life and consciousness are thus connected. Inference in single-celled organisms and the integration of information across the computational hierarchies of the human brain share a single principle: the incorporation of the state and causal architecture of the environment.

Furthermore, this blueprint explains how the abstraction we experience as consciousness can emerge from sustained resonant states described by the Adaptive Resonant Theory. Brain resonances may support conscious states by realizing information across multidimensional

mathematical objects. Such objects form and disintegrate as self-referential abstractions generated by iterations of local and global neuronal connectivity.

In contrast, I show how the most popular theory that addresses the Hard Problem, the Integrated Information Theory, is fundamentally flawed as an explanation for the cause of consciousness. It is based on a circular reasoning and, despite its qualities; it represents a cryptic form of intelligent design theory.

The ideas contained here occupied my mind as I went for dinner with a friend, who reminded me of a completely different angle to the same scientific search. I followed a thread that connected structures in single-celled organisms, to functional relationships, to informational relationships, to sensory systems, to globalization of information, to realization of information in abstractions, but he asked, “Have you ever thought that people might consider this discussion relevant to their spiritual pursuits?”

Over the years, when engaged in conversations about life after death or the possibility of a soul, I often heard skeptics challenge the spiritual stronghold by asking how a mere abstraction such as the soul could make a neuron fire.

My good friend pointed out that, by inverting the thread I had been following, I could actually answer this question. If the transition from physical to abstract goes both ways, then something that lacks any physical structure, the void can indeed make a neuron fire. I am not a religious man, and by no means would I attempt to stretch the reach of this work to the realm of religion. Nonetheless, the system I propose here allows for the generation of abstractions from the concrete, as it allows for the interference of the abstract on the biology of the human brain. I decided to honor and recognize my friend's observation.

In summary, through the course of this book, I outline a blueprint for the solution of one of the most difficult problems in neuroscience and philosophy, the Hard Problem of Consciousness. What started as a narrow question may appeal to a larger audience. Nevertheless, I left my considerations about abstractions as the enablers of physical changes in neurons to an appropriately brief last section, almost as a post hoc. At that point the reader will most likely have made his or her mind about the whole scheme.

Naturally, in studying such a complex problem, I often found myself outside my area of expertise. The neuroscience summarized in this book heavily relies on the work and opinion of major scientists in each different area, who have the credit for all the precise neurobiological depictions. I am fully responsible for any inaccurate interpretation of their perspectives. As I leaned, directly and indirectly on the work of other researchers, I take credit only for the application of their ideas and concepts to test the fitness of the hypotheses advanced by this book.

This book is divided in five sections. I review the Hard Problem of Consciousness in the first section, which includes the shortcomings of current theories, as well as a discussion of the computational and non-computational aspects of the problem. In particular, I explain why the Integrated Information Theory stems from a logical fallacy.

I detail the blueprint in the second section: a variation of Hofstadter’s strange loop that uses

potential states to explain how something can be generated out of nothing. This section also shows the reason why, from a computational perspective, “embodied abstractions” provide the best representations of conscious experiences. Finally, still in section two, I explain how life and consciousness share the same principles of statistical physics.

I test these hypotheses in the third and fourth sections of the book. I wrote them with no preconceived ideas, other than an intuition about how “processual expectations” underlie the kind of information eventually manifested as human consciousness. I studied local networks in search of the moment when they would become capable of hosting focal aspects of awareness. I was curious about the “why”.

I addressed the baffling thought of focal awareness: the brain is empty of consciousness much like local networks and single-celled organisms. The experience of being there that emerges from human consciousness is an illusion of the void. Nevertheless, I address the tension between focal awareness and the global sharing of information that characterizes human consciousness in section four. This section ends with an explanation of how consciousness may emerge even in the absence of the cerebral cortex.

The last section includes the most important thought experiment of this book. The “problem of records” reduces the Hard Problem to a specific set of questions and answers. The “problem of records” is not algorithmically compressible. Thus, it has no computational solution. It necessarily demands a physical, rather than a metaphysical answer.

I believe the “problem of records” may eventually allow for a serendipitous solution to the Hard Problem of consciousness. Such a solution may emerge from the symmetry between equations describing abstract “objects” that emerge from cross coupling of neuronal oscillations and the mathematical description of natural phenomena. This section indicates how to connect computational theories of consciousness with theories based on quantum mechanics.

This book can be read in different orders. Readers with a background in biology may wish to start from the sections focused on neuroscience. However, they may find some of the analysis perplexing. For example, it may seem that I treat focal awareness and global consciousness as something concrete. In the initial sections, readers will notice that I do exactly the opposite. I created a variable to depict the elusive “traces” of consciousness across hierarchical levels. The abstractness of the variable allows for the kind of treatment I later apply in the neuroscience review.

Those without a background in biology may find it easier to follow the order of the chapters and start with the philosophical aspects and the review of theories. The discussion starts at an abstract level but becomes practical and applied throughout the neuroscience review.

I believe that is likely easier to start from the chapter entitled Final Conclusions, then go through each chapter summary, and finally move to the beginning. This is optimally followed by a new look at the text with emphasis on footnotes. Many footnotes expand on the text and provide clarifying information.

I find the use of terms in the literature confusing and overlapping. In this book, I mostly use the word “awareness” for specific aspects of consciousness, such as awareness of location. I

do not use the word “consciousness” to address mental states such as the loss of consciousness during a coma. Here, the word refers to phenomenal consciousness, such as the consciousness of objects, colors, and sounds in an immediate and non-reflexive manner. Higher, global experiences, including the narrative consciousness and self-awareness of human beings, are thus preferentially addressed by the word “consciousness” instead of “awareness”.

I would like to thank Dr. Leonardo Ganem for reviewing the manuscript. I am grateful for his advice on science and language. I would like to thank those friends who endured our long and often repetitive conversations about consciousness – even when they had no knowledge of the subject, their support gave me confidence to continue on this journey. You all know who you are. Finally, I would like to thank the reviewers of this manuscript. I do not know who you are, but I thank you for the helpful feedback. I hope the good qualities of this book outweigh its shortcomings.

The subject is void, and the mind its avatar.

Paulo J. Negro, MD PhD
Assistant Clinical Professor of Psychiatry and Behavioral Sciences at the George Washington School of Medicine and Health Sciences
Assistant Professor in the Department of Psychiatry and Behavioral Sciences at Howard University
Chief Medical Officer
Kolmac Outpatient Recovery Centers

ACKNOWLEDGEMENTS

Ideas have pedigree. I would be remiss to not recognize those whose work greatly influenced this publication. György Buzsáki for all things oscillatory. Douglas Hofstadter for the strange loop. Terrence Deacon for the incompleteness of things, Dana Ballard for the hierarchy of abstractions. And Ezequiel Morsella, not only for the action synthesis, but for his kindness. There are many others that influenced this book, as recognized in the text. I thank you all. I should extend my gratitude to Ginger Campbell whose podcast awakened old thoughts and feelings about the origin of human consciousness. And, finally, to Kōun Yamada Roshi whose Gateless Gate still sits on my nightstand.

NOTE

[2]In Italian: “Io ho capito” (I realized), in Portuguese: “Eu compreendi” (I realized or I understood).