Introduction

At the invitation of Charles Smith and Carl York of the System Development Foundation, I began in 1985 to plan a symposium whose purpose was to demonstrate the status of current attempts to understand the nature of neural computation and the interaction of computationally motivated research in the brain sciences. Letters were sent in July of 1985 asking for authors to contribute to a book whose purpose was to define the term *computational neuroscience*. The chapters of this book were presented for group discussion at a symposium held in June 1987 at Carmel Highlands, California.

Computational Neuroscience

The System Development Foundation provided an excellent opportunity to explore the fundamental structure of a field in which I have worked for over fifteen years, but one whose fortunes, and even name, have experienced violent turbulence over the past four decades. Over the years, the terms *cybernetics*, *neural networks*, *brain theory*, *artificial intelligence*, and *neural modeling* (among others) have been applied to this general area. Each term has been used and abused, and none felt untainted for the present purpose. Since terms like *computational fluid mechanics* have been in use for some time without any apparent overloading, the term *computational neuroscience* was chosen. One advantage of this term is that there did seem to be several readily available role models to help provide the definition to go along with the terminology. Fluid mechanics is a field in its own right, but computational fluid mechanics is a specialty in which there is sufficient overlap between computer science and fluid mechanics to justify adding new nomenclature. Computational neuroscience might be characterized as that area of overlap between neuroscience and computer science which required sufficient specialized expertise to justify a new subdiscipline.

The defining characteristic of computational neuroscience, then, is the problem area in which *difficult algorithmic or implementational questions are intimately related to the data of the nervous system.* The interplay of neural data and of computation and applied mathematics define the scope of this term.

Not every application that claims to model a "neural network" is computational neuroscience: if a particular application has no grounding or constraint in actually observed neural data, as seems to be the case for much of the current "neural network" literature, then it would be less confusing to use some other term. Conversely, not every research application that models neural data with the help of a computer should be called computational neuroscience. Not every neuroscientist using a computer is a computational neuroscientist! Once again, we can fall back on our previous metaphor: clearly, not every physicist who uses a computer is called a computational physicist. The pragmatic motivation for being careful about these details is to avoid the exaggeration, confusion, and overheated salesmanship that have caused the previous incarnations of this field to collapse periodically.

The contributors to this book were invited to present their own definitions of computational neuroscience, and the first part of the book provides several alternative definitions. One of these contributions, in revised form, has been republished (Churchland, Koch, and Sejnowski, *Science* 240 (1989)), and has begun to popularize this terminology. With some luck, *computational neuroscience* will provide a reasonably precise and clean term to describe an area of research whose time has finally come.

Plan

The initial plan for this book was worked out in a series of discussions with Yehezkel Yeshurun, and was further refined in a meeting with a steering committee consisting of David Zipser, Shimon Ullman, and Terry Sejnowski in Cambridge, Massachusetts, on November 7, 1985. The plan was to solicit chapters from several dozen colleagues which would illustrate by example the scope of current work in computational neuroscience. This work was organized by physical scale: synaptic, neuronal, topographic/columnar, and "systems." In addition, authors were encouraged to contribute "philosophical or historical" chapters and discussions of the term *computational neuroscience*. This provided an unusual opportunity for examination of the fundamental and structural aspects of this discipline. Rall's chapter on his reminiscences of the early days of the field, Daugman's analysis of the history of technological metaphors for the brain, Grobstein's discussion of the applicability of lesion technique, Grossberg's analysis of Liapunov approaches to neural modeling, and discussions of the scope of computational neuroscience by the late Donald Perkel and by Churchland, Koch, and Sejnowski make up Part I of the book. The remaining parts follow the ascension in scale from the synaptic, neuronal, map, and system levels, and provide a cross-section of contemporary research in computational neuroscience.

Structural Instability of Computational Neuroscience

One additional aspect of this project deserves mention. It is often observed that the general subject area of this book has seemed to undergo a historical cycle of boom and bust. We are currently in the midst of a boom, characterized by intense interest in the area of "neural networks." One characteristic feature of this boom phase is captured by the following quote:

There seem to have been [several] main reasons for negative reactions to [neural networks]: first was the admitted lack of mathematical rigor in preliminary reports. Second was the handling of the first public announcements ... by the popular press, which fell to the task with all of the exuberance and sense of discretion of a pack of happy bloodhounds. Such headlines as "Frankenstein Monster Designed by Navy Robot That Thinks" (Tulsa Times*) were hardly designed to inspire scientific confidence.*

This quote, which sounds familiar and contemporary, is actually from Frank Rosenblatt's 1962 book *Principles of Neurodynamics*. Coming at the height of an earlier boom phase of interest in neural computation, the early perceptron movement faded away in the late 1960s, only to revive again in this decade with the current "neural network" movement.

Why is it that one of the last major scientific frontiers, the attempt to understand the basic principles of neural computation, is so structurally unstable? In an attempt to understand this phenomenon, and hopefully to avoid its recurrence, it was suggested to the steering committee that it would be useful to study the fundamental questions that needed to be answered in computational neuroscience. Partial motivation for this was provided by the set of questions that Hilbert formulated, at the turn of the century, in an attempt to frame the basic unanswered questions in mathematics.

It became clear almost immediately that this was not mathematics and that we were not Hilbert. Proceeding to a less ambitious but more appropriate quest, it was suggested that we try to frame at least some of the basic unanswered questions in computational neuroscience. One of the problems in this field is that there is relatively little consensus about the basic facts and important questions. This lack of consensus leads unavoidably to confusion; to the extent that it is largely unacknowledged, and subconscious, it could be one of the causes of the instability of the field. It seemed that we could provide an important service by sending a list of some basic controversial issues to the authors of the book for further comment and discussion. A dialectic form was chosen: a given position was presented in two opposing statements addressing the same issue. These dialectics were viewed as "conceptual singularities," in the sense that the term *singularity* is used to describe a place where structure breaks down and where no well-defined statement can be made. The hope was that, by explicitly stating 5 or 10 such singularities, we would highlight the underlying conceptual problems in computational neuroscience and provide a well-defined (and highly charged) platform for further discussion.

An organizational meeting was held in San Diego on June 12 and 13, 1986. This meeting was attended by Michael Arbib, Carver Mead, Donald H. Perkel, Terry Sejnowski, Eric Schwartz, Yehezkel Yeshurun, and David Zipser. Roberta Ishihara, Charles Smith, and Carl York attended as representatives of the System Development Foundation. This meeting was great fun and highly spirited. The term *singularity* seemed to be well chosen, since the participants had to be frequently reminded that our job was to frame these dialectics, not to adopt a position within them! From a prototype set of eight dialectics which was presented to this group, a final set of seven was framed, which is reproduced below.

The original intention of discussing these issues in depth at the Symposium on Computational Neuroscience, which was held in Carmel on June 5, 1987, never came about. Interestingly, the participants at this meeting, who included the authors of the chapters in the present book as well as additional observers, split into two camps vis-à-vis the desirability of this discussion. About half of the participants were eager for, and half averse to, discussion of these fundamental dialectics. Obviously, this requires the addition of an additional dialectic: "It is, or is not, a waste of time to discuss fundamental structural issues of computational neuroscience."

At the end of the Carmel symposium there was demand to have a mini-symposium in order to pursue this discussion. That meeting was attended by Marc Jeannerod, Don Glaser, Scott Weinstein, Carver Mead, John Daugman, Carl York, Yehezkel Yeshurun, Eric Schwartz, Alan Rojer, Cathryn Downing, Daphna Weinshall, and Terry Sejnowski. One of the requests that came out of the mini-symposium was to present the list of dialectics in the present book, since they represented a stimulating set of questions which were likely to be of interest to at least some of the readers.

The formulation of this list was greatly assisted by the colleagues listed above. To the extent that this activity has any merit at all, it is owed to the ideas and work of the participants. It cannot be too strongly stressed, however, that the final wording and presentation of these issues was done by myself, and I take sole responsibility for any errors of wording, concept, or presentation.

The following list will probably have little utility to sophisticated workers in computational neuroscience, since the positions which are presented dialectically have roughly equal validity. They are meant to illustrate an underlying unresolved issue, not to provide material for debate! But if the reader finds himself identifying strongly with one or the other of these contrasting positions, or is surprised to find that one of the alternate positions is seriously proposed, then perhaps the efforts expended in framing this list will not have been wasted.

Some Basic Dialectics of Computational Neuroscience

1. What is the relevance of simple model neural systems?

Position 1: Mammalian nervous systems are far too complex to study the properties of neural networks. We must begin with the simplest model system available, either invertebrate organisms or simple localized synaptic modules in vertebrates.

Position 2: Two decades have gone by, and the early rhetoric of model-systems proponents has not been fulfilled. Brute-force analysis of simple systems will not lead anywhere. Invertebrate behavioral capabilities and neural properties are remote from those of vertebrates. Simple invertebrate systems are of interest only in themselves, not as general brain models. Synaptic-level studies have not provided much insight into computational function.

2. Is physical locality an essential part of neural computation?

Position 1 (homage to the Turing machine): The anatomical structure of the brain has no more to do with its function than the shape of the cabinet of a VAX, or the location of its circuit boards. Brain function is determined by the logical and dynamic connection properties of its neurons. The actual physical structure, location, architecture, and geometry is irrelevant compared to its logical, connectionist aspects. One could take a brain and grossly deform the positions of its neurons, keeping only the topology of connections intact, and there would be negligible difference in performance.

Position 2 (computational anatomy): Significant recent work in brain research has been related to the discovery and elucidation of detailed forms of somatotopic mapping, laminar specialization in cortex, and columnar architectures representing sensory submodalities. These forms of functional architecture may represent a major mode of brain function: the formatting of sensory data in a manner that simplifies its further processing. One of the major differences in computational style of brain versus VAX may well be the indifference of the VAX to its geometry and the exquisite attention paid by the brain to its geometry.

3. What is the appropriate balance of theory and experiment in neuroscience?

Position 1: Theoreticians have constructed models which have little connection to experimental data, and which provide little opportunity for experimental test. Experimentalists will learn whatever is necessary to perform their work. So far, they have not had to bother with learning theory (mathematics, computer science, . . .). Theoreticians will have to do their homework, i.e. learn something about the experimental disciplines which they are modeling.

Position 2: Experimentalists have too little background to appreciate or understand the relevance of theoretical work. A science with no theoretical component is just a mass of phenomenological details, a form of "butterfly collecting." Experimentalists will have to do their homework, i.e. learn something about the theoretical tools which are essential to their disciplines.

4. What is the correct choice of spatial scale for modeling the computational abilities of brains?

Position 1 (synaptic-neuronal level): Since the brain is composed of neurons, whose individual properties are accessible to study and fairly well understood, the neural scale is the correct one at which to approach brain computation. Brains are networks of neurons, and the mathematical properties of such networks determine brain function.

Position 2 (column-map level): The mathematical tools available for studying large networks are woefully inadequate. The two-body problem in mechanics is easy; the three-body problem is very hard. The neural *N*-body problem presents difficulties vastly greater than those of classical statistical mechanics, a field which is largely computationally intractable. Neural firing densities need to be averaged into "densities," and the large-scale (i.e. columnar, map) properties of these densities made the basis for study. Just as fluid mechanics began with a simplified continuum hypothesis, neural modeling must find a simplified continuum level in order to "get off the ground."

5. What is the correct choice of temporal scale for modeling neural computation?

Position 1: The brain may be viewed as a dynamical system. Stable states of neural networks may be reached, since the time constant of neurons is in the range of 1–10 milliseconds, and differential equation systems have been shown which reach their stable states in only a few iterations.

Position 2: Pre-attentive perception refers to the period about 200 milliseconds after an event. Most perceptual and many cognitive functions can occur in roughly this time period. But 200 milliseconds is also roughly the time of transit for a signal through the brain. It appears that the brain is a one-cycle machine, something like a lookup table! There is no time for any settling into stable states. Moreover, there is no experimental evidence that the brain ever "settles" into anything like an equilibrium condition.

6. Is the contemporary Pavlovian position correct?

Position 1: Perception cannot be studied in isolation from concomitant motor or other goal-directed activity.

Position 2: One can show, for example, a Julesz random-dot stereogram to someone who has never seen one before, and this person will see a stereo percept, without even knowing what he is looking at. Perception can be isolated, both functionally and computationally, from goal-directed activity.

7. Are conventional computer metaphors valid for the brain?

Position 1: It is crucial to distinguish algorithm from implementation in AI and neural modeling (David Marr). The algorithm need have nothing to do with the underlying physical structure of the neuronal "hardware."

Position 2: The implementation (brain structure) depends crucially on the nature of the computation (brain function). In the brain, the medium is the message.

Acknowledgments

This book, and the SDF Symposium on Computational Neuroscience at Carmel Highlands, were made possible by generous funding from the System Development Foundation, and by the Life Sciences Directorate of the Air Force Office of Scientific Research. Special thanks are due to Charles Smith and the late Augustus Kinsel III (SDF) and John Tangney (AFOSR) for their early and enthusiastic support for research in the area of computational neuroscience.

The planning of this book and symposium owe much to early discussions with Yehezkel Yeshurun, and to the suggestions of the steering committee of Terry Sejnowsky, Shimon Ulman, and David Zipser. Thanks to Alan Shaw for his editorial assistance, and to Alan Rojer for many discussions.

I dedicate this book to the memory of my father, Jack Schwartz, whose life-long interest in science and mathematics inspired my own, and to my wife, Helen, who originally sparked my interest in studying neuroscience and who nurtured the change from high-energy physics which ultimately led to the present work.

Contents