TWO SCIENCES OF MIND

ADVANCES IN CONSCIOUSNESS RESEARCH

ADVANCES IN CONSCIOUSNESS RESEARCH provides a forum for scholars from different scientific disciplines and fields of knowledge who study consciousness in its multifaceted aspects. Thus the Series will include (but not be limited to) the various areas of cognitive science, including cognitive psychology, linguistics, brain science and philosophy. The orientation of the Series is toward developing new interdisciplinary and integrative approaches for the investigation, description and theory of consciousness, as well as the practical consequences of this research for the individual and society.

Volume 9

Seán Ó Nualláin, Paul Mc Kevitt and Eoghan Mac Aogáin (eds)

Two Sciences of Mind
Readings in cognitive science and consciousness

TWO SCIENCES OF MIND

READINGS IN COGNITIVE SCIENCE AND CONSCIOUSNESS

Edited by

SEÁN Ó NUALLÁIN
National Research Council, Ottawa
Dublin City University

PAUL MC KEVITT
Aalborg University

EOGHAN MAC AOGÁIN
Linguistics Institute of Ireland, Dublin

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Two sciences of mind : readings in cognitive science and consciousness / edited by Seán Ó Nualláin, Paul Mc Kevitt, Eoghan Mac Aogáin.
p. cm. -- (Advances in consciousness research, ISSN 1381-589X ; v. 9)
Papers originally presented at a workshop on "Reaching for Mind."
Includes bibliographical references.
1. Cognitive science--Congresses. 2. Consciousness--Congresses. I. Ó Nualláin, Seán. II. Mc Kevitt, Paul. III. Mac Aogáin, Eoghan. IV. Series.
BF311.T87 1997
153--dc21 96-52164
ISBN 90 272 5129 0 (Eur.) / 1-55619-189-8 (US) (Pb; alk. paper) CIP

John Benjamins Publishing Co. • P.O.Box 75577 • 1070 AN Amsterdam • The Netherlands
John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

Table of Contents

About the Editors

Seán Ó Nualláin holds an M.Sc. in Psychology from University College and a Ph.D. in Computer Science from Trinity College, both in Dublin, Ireland. He has just returned from sabbatical leave at the National Research Council (NRC), Canada to his lecturing post at Dublin City University, Ireland where he initiated and directed the B.Sc. in Applied Computational Linguistics. He is the author of a book on the foundations of Cognitive Science: *The Search for Mind* (Ablex, 1995). He chaired the first workshop on which this book is based (co-chair Paul Mc Kevitt) at Sheffield in April, 1995.

Paul Mc Kevitt is a Visiting Professor of Intelligent Multimedia Computing at Aalborg University in Denmark, and a British EPSRC (Engineering and Physical Sciences Research Council) Advanced Fellow in the Department of Computer Science at the University of Sheffield, England. The Fellowship releases him from his Associate Professorship for five years to conduct full-time research on the integration of natural language, speech and vision processing. He is currently pursuing a Master's Degree in Education at the University of Sheffield. He completed his Ph.D. in Computer Science at the University of Exeter, England in 1991. His Master's Degree in Computer Science was obtained from New Mexico State University in 1988 and his Bachelor's Degree in Computer Science from University College Dublin in 1985. His primary research interests are in Natural Language Processing including the processing of pragmatics, beliefs and intentions in dialogue. He is also interested in philosophy, Multimedia and the general area of Artificial Intelligence.

Eoghan Mac Aogáin is Director of Instidiud Teangeolaiochta (Linguistics Institute) of Ireland. He studied philosophy at University College Dublin, Ireland and psychology at the Centre for Advanced Study in Theoretical Psychology at the University of Alberta, Canada. He has participated in a number of EU Natural Language Processing (NLP) projects, and is currently involved in PAROLE, which will provide computational lexica and corpora for 13 Union languages to a common standard. His interests include corpus-based approaches to grammar writing and the design of language learning programs. He is the founder editor of the international journal: Language, Culture and Curriculum.

List of Contributors

Kenneth Aizawa
Dept. of Philosophy
Centenary College
2911 Centenary Blvd
SHREVEPORT, LA 71134-1188, USA
kaizawa@beta.centenary.edu

Bernard J. Baars
The Wright Institute
2728 Durant Avenue
BERKELEY CA 94704, USA
baars@cogsci.berkeley.edu

John A. Barnden
Computing Research Lab &
Comp. Science Dept.
New Mexico State Univ.
LAS CRUCES, NM 88003-8001, USA
jbarnden@crl.nmsu.edu

István S.N. Berkeley
Dept. of Philosophy
Univ. of Southwestern
Louisiana, P.O. Box 43770
LAFAYETTE, LA 705-3770, USA
istvan@usl.edu

Mark H. Bickhard
Dept. of Psychology
Lehigh University
17, Memorial Drive East
BETHLEHEM, PA 18015-3068, USA
mhb0@lehigh.edu

Andrew Brook
2217 Dunton Tower
Carleton University
1125 Col. By Drive
OTTAWA, ON K1S 5B6, Canada
abrook@ccs.carleton.ca

Sung-Bae Cho
Dept. of Computer Science
Yonsei University
SEOUL, Korea

Dr Thomas Christaller
German National Center for
Information Technology GMD
AI Research Division
Sloß Birlinghoven
D 53754 Sankt Augustin, Germany
christaller@gmd.de

Terry Dartnall
Griffith University
Fac. Science & Technology
Nathan Campus, Kessels Rd
NATHAN, BRISBANE QUE 4111,
Australia

Kerstin Dautenhahn
VUB-AI Lab.
Autonomous Agents Group
Bldg G10, Room 725
Pleinlaan 2
1050 BRUSSELS, Belgium
Belgiumkerstin@arti9.vub.ac.be

Craig DeLancey
Indiana University
Dept. Philosophy
Sycamore Hall 026
BLOOMINGTON IN 47405-2601, USA

Chenming Guo
Computers Science Dept.
Tsinghua University
BEIJING 100084, China
cmguo@sun.ihep.ac.cn

William C. Hoffmann
2591 W Camino Llano
TUSCON AZ 85742, USA

B.F. Katz
School of Cognitive and
Computing Sciences
University of Sussex
BRIGHTON BN1 9QH, UK

Phil Kime
Dept. of Philosophy
Univ. of Edinburg, David
Hume Tower, George Sq.
EDINBURGH EH8 9JX, UK
Phil.kime@ed.ac.uk

Eoghan Mac Aogáin
Inst. Teangeolaiochta
(Irish Linguistics Inst.)
31, Plás Mac Liam
Dublin 2, Ireland
Eoghan.MacAogain.M248@eurokom.ie

K.F. MacDorman
Computer Laboratory
Pembroke Street
CAMBRIDGE CB2 3QG, UK
karl.macdorman@cl.cam.ac.uk

Paul Mc Kevitt
Center for Pesonkommunikation
Fredrik Bajers Vej 7-A2
Institute of Electric Systems
Aalborg University
DK 9220 AALBORG, Denmark
pmck@cpk.auc.dk

Ajit Narayanan
University of Exeter
Dept. of Comp. Science
Old Library
EXETER EX4 4PT, UK
Ajit@dcs.exeter.ac.uk

Prof.dr. James Newman
Colorado Neurological
Institute
740 Clarkson Street
DENVER, CO 80806, USA
newmanjb@aol.com

Seán Ó Nualláin
Dublin City University
Baile Átha Cliath 9
EIRE/IRELAND
sean@CompApp.DCU.IE
sean@ai.iiT.nrc.ca

N.C. Riley
School of Cognitive and
Computing Sciences
University of Sussex
BRIGHTON BN1 9QH, UK

Gerard Sabah
Language and Cognition
Group (LIMSI), CNRS
B.P. 133
F 91403 ORSAY Cedex, France
gs@limsi.fr

Dr P.D. Scott
Dept. of Computer Science
University of Essex
COLCHESTER CO4 3SQ, UK
scott@essex.ac.uk

J.G. Taylor
Centre for Neural Netwoks
Dept. of Mathematics
Kings College, Strand
LONDON WC2R 2LS, UK
udah057@bay.cc.kcl.ac.uk

Stuart Watt
Department of Psychology
The Open University
MILTON KEYNES, MK7 6AA, UK
S.N.K.Watt.@open.ac.uk

Introduction

Seán Ó Nualláin
Dublin City University and NRC, Canada

The "Reaching for Mind": Foundations of Cognitive Science (CS) workshop was announced over the Internet as follows:

1. Workshop Description

The assumption underlying this workshop is that Cognitive Science (CS) is in crisis. The crisis manifests itself, as exemplified by the recent Buffalo summer institute, in a complete lack of consensus among even the biggest names in the field on whether CS has or indeed should have a clearly identifiable focus of study; the issue of identifying this focus is a separate and more difficult one. Though academic programs in CS have in general settled into a pattern compatible with classical computationalist CS (Pylyshyn 1984, Von Eckardt 1993), including the relegation from focal consideration of consciousness, affect and social factors, two fronts have been opened on this classical position.

The first front is well-publicized and highly visible. Both Searle (1992) and Edelman refuse to grant any special status to information-processing in explanation of mental process. In contrast, they argue, we should focus on Neuroscience on the one hand and Consciousness on the other. The other front is ultimately the more compelling one. It consists of those researchers from inside CS who are currently working on consciousness, affect and social factors and do not see any incompatibility between this research and their vision of CS, which is that of a Science of Mind.

2. Workshop Issues

The tension which riddles current CS can therefore be stated thus: CS, which gained its initial capital by adopting the computational metaphor, is being constrained by this metaphor as it attempts to become an encompassing Science of Mind. Papers are invited for this workshop which:

1. Address this central tension.
2. Propose an overall framework for CS (as attempted, *inter alia*, by Ó Nualláin (1995)).
3. Explicate the relations between the disciplines which comprise CS.
4. Relate educational experiences in the field.
5. Describe research outside the framework of classical computationalist CS in the context of an alternative framework.
6. Promotes a single logico-mathematical formalism as a theory of Mind (as attempted by Harmony theory and using category theory).
7. Moderately or indeed violently disagree with the premise of the workshop.

Ó Nualláin, S. 1995. *The Search for Mind: A New Foundation for CS.* Norwood: Ablex.
Pylyshyn, Z. 1984. *Computation and Cognition.* Cambridge, MA: MIT Press.
Searle, J. 1992. *The Rediscovery of the Mind.* Cambridge, MA: MIT Press.
Von Eckardt 1993. *What is Cognitive Science?* Cambridge, MA: MIT Press.

The Workshop Committee was as follows:

John Barnden	(New Mexico State University, NM, USA & University of Reading, England)
Istvan Berkeley	(University of Alberta, Canada)
Mike Brady	(Oxford, England)
Harry Bunt	(ITK, Tilburg, The Netherlands)
Daniel Dennett	(Tufts University, USA)
Eric Dietrich	(SUNY Binghamton, NY, USA)
Jerry Feldman	(ICSI, UC Berkeley, USA)
Stevan Harnad	(University of Southampton, England)
James Martin	(University of Colorado at Boulder, CO, USA)
Eoghan MacAogain	(Instidiud Teangeolaiochta/Irish Linguistics Institute, Dublin, Ireland)
John Macnamara†	(McGill University, Canada)
Mike McTear	(Universities of Ulster and Koblenz, Germany)
Ryuichi Oka	(RWC P, Tsukuba, Japan)
Jordan Pollack	(Ohio State University, OH, USA)

Zenon Pylyshyn	(Rutgers University, USA)
Ronan Reilly	(University College, Dublin)
Roger Schank	(ILS, Northwestern, Illinois, USA)
Walther v.Hahn	(University of Hamburg, Germany)
Yorick Wilks	(University of Sheffield, England)

† We regret to say that John Macnamara passed away in January, 1996

A remarkable 40 papers had been submitted within two months of the announcement. The fact that so many distinguished members of the CS community agreed to act as members of a committee discussing a crisis in the foundations of their discipline also was telling. The workshop itself was an extremely lively affair, as may be inferred from the diversity of approaches manifest in the papers in each of the three separate parts of this book. The final discussion focussed on two issues:

- the relation between epistemology and ontology,
- what is semantics?

The latter issue recurs throughout this book; the former is implicit in, inter alia, the debate on what precisely to do with consciousness. I have found myself forced to conclude that it is necessary to found a separate science of consciousness, combining normal "science" with phenomenal analysis, alongside a CS based (less controversially) on information. Hopefully, after reading part 3, readers will be likewise convinced.

Part 1 ("CS in Crisis? Cognition and Mind") features papers mainly addressing workshop issues 1, 2 and 7; Part 2 concerns itself more with 4. In Part 3, we will encounter a series of *Weltanschauungen* (all-encompassing views) of great consequentiality which confront all the workshop issues, and more. I hope this book will convey some of the excitement of the event. Some papers which are accepted by the committee but not presented at the workshop due to valid reasons also are included.

3. Acknowledgments

As often happens, Paul Mc Kevitt woke me from my slumbers to suggest I give a public expression to what had been half-formulated thoughts by writing the workshop spec. His work as publicist and general scrummager was invaluable. Eoghan Mac Aogáin's contribution as a reviewer was such that this book would not otherwise have been produced. I wish to thank those who provided extra reviews, in particular the Canadian trio of Arnold Smith, Rob Stainton, Peter

Turney. Any errors that may remain are my responsibility, and remain there even after Noel Evan's best efforts to rein me in. Finally, I dedicate this book to all researchers and teachers in this area who are open-mindedly trying to do justice to the bewildering range of data obtained.

Seán Ó Nualláin

Part I: Cognitive Science in Crisis?

Cognition and Mind

Seán Ó Nualláin

Phil Kime's lucidly-written paper provides an excellent introduction to this part. Kime agrees with the premise of the workshop, but for reasons other than those given in its description. His argument is two-pronged. First of all, the "semantic" formalisms used in CS, particularly where its concerns converge with those of linguistics, are borrowings from the work of logicians like Tarski and Frege who warned of their inappropriateness for natural language. Secondly, different CS camps have wildly different notions of evidence and explanation.

One particular issue straddles the two pillars of the argument; the relation between formalism and *datum.* Since they normally are *a priori* constructions which are not built on any systematic correspondence with the data, there is no difficulty in extending formalisms to handle any given new *datum.* The force of their claim does not diminish in this case because it is essentially an artifact of their internal logical consistency, which can be preserved. Much of cognition is, Kime continues (in logical terms) contingent rather than necessary and will be captured, if at all, by non-Euclidean geometric systems and the like. Indeed, these might yet supply a counter-argument to the systematicity objection (see Aizawa's paper in Part 2). Almost as a parting shot, Kime takes the sacred cow of native speaker intuitions to task with respect to the evidential and explanatory schisms he sees.

Narayanan agrees with the premise of the workshop. He begins by citing some objections to classical CS and AI, i.e., rule-following is not enough for intelligent behavior and in any case falls prey to Gödelian arguments. However, his major concerns are other; he claims that even the eliminative materialists' bet that cognition will eventually be explained with respect to neural activity is too sure to be interesting. He wishes to examine cognition at a finer level of granularity, i.e. that of biochemistry. Emboldened by the example of Penrose, who with Hameroff suggests that the cytoskeleton of cells will provide answers about the seat of consciousness, Narayanan takes us on a biological tour. There is an implication that CS is in some way scientifically obliged to ground its descriptions at this level. Granted, DNA has been already speculated about as a computational device (e.g. solving the travelling salesman's problem) and protein-folding may in general be computationally exploited. However, several good reasons exist for not journeying in this particular direction.

The first is, quite simply, that reductionism is insatiable. Having rejected the Churchlands' granularity as too coarse, we are logically obliged to end at the subatomic level. Secondly, we lack biomolecular explanatory mechanisms for any aspects of cognition. Thirdly, if as Watt later argues, eliminative materialism gives free rein to the mind to settle on (perhaps outlandish) metaphors to understand itself, biomolecular CS will encourage even more promiscuous behavior.

Unsurprisingly, my paper addresses itself to practically all of the workshop issues, with varying degrees of coverage. The notion of "crisis" is interpreted as "opportunity" as well as "quagmire." Von Eckardt's recent characterization of CS is used to focus on the tension that inevitably arises when an essentially informational notion of cognition confronts the phenomenon of consciousness. Several positive recommendations are issued; the domain of CS is to be those aspects of mind which can be informationally characterized, with obvious consequences for its hitherto nebulous academic domain. Several synthetic themes arising from the disciplines which inform CS are then discussed, and the sources of evidence for "egocentric" cognition proposed as a proof of concept for argument from synthesis. Finally, it is argued that a neuroscientific "invasion" may be salutary for, rather than destructive of CS.

For the next two writers CS is not in crisis. Watt argues that metaphors are the stuff of science; somewhat controversially, he states that even the "substantive assumptions" of CS as described by Von Eckardt are metaphorical. We cannot but be anthropomorphic; we project ourselves onto every phenomenon, and Searle's Chinese Room argument gains its force from this. (Barnden extends this type of reasoning in Part 3). Scott's approach is much harder-nosed; scientific theories are communicable and thus inevitably Turing-computable. Cognitivism, the notion that mentation is computation, must be scientifically correct; any theory of mind is essentially a program. However, there are two flaws in Scott's argument. The first is that communication can occur between initiates on matters (like art) for which it is excessive to predicate "computability" as he wishes to use the term. Secondly, his treatment of consciousness is quite simply incorrect; as the papers in Part 3 demonstrate, there is a consensus that it has a functional purpose. However, Scott's paper alerts us to the necessity of finding ways to handle social factors, emotion and consciousness informationally and our inability to handle subjectivity.

Dartnall's substantial contribution revives an old issue. Our minds somehow seem to cope with logical systems employing notions like necessity and certainty. There are two fallacious explanations for this, each flawed in its way. The first, logicism (or reverse psychologism) insists that the laws of logic are the laws of thought. Recent culprits, Dartnall argues, include Noam Chomsky and John MacNamara. The second fallacy, psychologism, attempts to reduce all acts of logical inference to purely psychological processes, which cannot have the characteristic of logical necessity.

Dartnall follows Kime's anti-logicism with a broad historical range of reference. He argues that the central dichotomy is "act versus content" i.e. confusion arises only when we fail to distinguish psychological acts and their intentional objects. The PSSH (physical symbols hypothesis) in AI tries to pack in both act and content; this is precisely the source of the Chinese Room argument's force. An appropriate alternative contrast, Dartnall concludes, is state versus content.

Finally, Bickhard eases the transition into the next part with a sustained attack on the standard story about representation. He argues that our current notion of cognition is parasitic on this story; its putative nature as an autonomous subsystem is equally misconceived.

Representation, (R) as classically conceived, involves "encodingism"; the external world is encoded in some form, which contains a Cartesian flaw. Lacking a homunculus, we need to ground representations otherwise. Some solace may be found in the Piagetian notion that R is an internalization of interactions midway between subject and object. Drescher has extended this work but a better model is one in which representation, action and motivation are seen as manifestations of a single underlying ontology.

It is indeed true that a set of concepts (like egocentric/intersubjective) may have to be superimposed on our notion of R; what is however more urgent for CS is an explicit realist stance. Bickhard points out the problems with R with considerable skill; my own guess is that the notion of context and the role of self in delimiting context will eventually be seen as crucial. However, this leaves us with the problem of how some of our most abstract constructions (like Riemann geometry) refer to anything; what Wigner encapsulates as "the unreasonable effectiveness of mathematics."

Reinventing the Square Wheel

The Nature of the Crisis in Cognitive Science

Phil Kime
Centre for Cognitive Science
University of Edinburgh

1. Preliminaries

Before bemoaning a crisis, it is always best to look around to see if there really is one. Ever since Montague, the momentum in semantics carrying it towards a formal theory of meaning for natural languages has been a significant spoke in the wheel of Cognitive Science. We learnt how to deal with the intensional identity problems that Montague left us with; we learnt how to construct detailed and computationally tractable models of compositionality and how to approach the incorporation of notions of tense, modality and quantification. For all the progress that the field boasts, you would hardly think there was a crisis at all.

However, this is really an illusion. Semantics and traditional theories of meaning are not in a good way and this is, in my opinion, central to the problem that besets Cognitive Science as it underlies the problem with the entire computational program. The crisis as I see it has two main threads: the first being that the computationalist program is wedded to formalisms and ideas that were imported wholesale from people who had principled reasons not to lend them out. The nature of the tools employed within the field are such that they restrict the natural and desirable criticism that a field must support. The current tension within Cognitive Science is, I think, partly a result of the dogma engendered by the formalist thrust one finds at its most stark in semantics. Secondly, there is quite a deep rift between the notions of evidence and explanation that different camps within Cognitive Science employ. Progress in a field is difficult when there is no general agreement about the sources and types of evidence used to test theories.

2. Origins of the Crisis

Primarily, the grip the computational program has on the field as a whole is a result of the misappropriation of logical formalism. In the workshops and everyday seminars on semantics, one hears the words "Fregean" and "Tarskian" mentioned with regularity. If one is to consult the works of the aforementioned, one finds a striking overall repellence to the regimentation of their formal creations as tools in the analysis of natural languages: something the computationalist program has adopted on a large scale. There was a general conception espoused by these progenitors of modern formal theory that natural languages were not sufficiently well defined and exact to allow treatment by formalisms that presupposed a certain amount of regularity and structure in their subject matter. For example, Frege says:

> Language is not governed by logical laws in such a way that mere adherence to grammar would guarantee the formal correctness of thought processes.[1]

So, here we see a concern, not for natural language, but with the thoughts lying behind it. Frege was skeptical about the application of his formalism to natural languages. While Frege was sometimes less than clear on this point, particularly during his early work, he makes quite strident remarks about the applicability of his formalism to natural language in places: more markedly in later work. However, we find this concern explicit in Tarski:

> … the concept of truth (as well as other semantical concepts) when applied to colloquial language in conjunction with the normal laws of logic leads inevitably to confusions and contradictions.[2]

Tarski was of the opinion that the application of the formal methods to natural languages would necessitate a reform of the language; hardly something that an explicative semantics would aspire to. After all, the reform of a natural language results in an artificial one, thus defeating the object an explanatory enterprise.

Now, formalists generally either ignored these warnings or sought to prove them groundless by devious formal innovations designed to appropriate, in the spirit of Davidson's famous paper on "Truth and Meaning," more and more features of natural language for formal description. Around this time a few were beginning to worry about the assumptions inherent in the formal methods advocated by this approach. For example, Hubert Dreyfus published his famous book on the shortcomings of traditional approaches in AI in 1972. This pointed to the lack of progress overall and suggested certain underlying assumptions were to blame. It seems that the current crisis in Cognitive Science and in particular in semantics is of the same sort: an underlying inadequacy of assumptions inherited from formalisms unsuited to the task; but today it is even worse.

What makes things worse is that we have learnt some new tricks to prevent overt crisis; some new tricks that make everything seem alright. If one looks at the sorts of stock examples that formalist semanticists deal with today, it is quite disturbing to note how simple they still are given the supposed applicability of the formalisms employed. Disturbing too to note how similar they are to examples used ten or twenty years ago. The reason is that we have mastered the art of getting fatter instead of getting further. By that, I mean that a problem is something that results, not in a reevaluation of the foundations of a theory, but almost exclusively in revision of technical minutiae. If we cannot deal with a particular example, we tweak the formalism until we can. More dramatically, we invent another formalism specifically designed to deal with the problem. Progress is seen to be the accommodation of errant data with little respect for the implications for the assumptions of a theory. The crisis in Cognitive Science amounts to exactly this. Problems have come to be dealt with entirely within the scope of the dominant formalist research program. If your formalism starts to give you problems with the representation of the meaning of a certain sentence, make a new formalism. I have lost track in recent years of the number of different semantic formalisms. We now have Dynamic Predicate Logic, Dynamic Montague Grammar, Discourse Representation Theory, Situation Theory, Property Theory, Channel Theory, Linear Logic and many more. Many of these were explicitly motivated by problems with a particular natural language construction. I recall attending a workshop dedicated to formal semantics research just a few years ago where a prominent linguistic semanticist was challenged "But isn't this new theory just tackling the same examples as previous theories have been tackling for over ten years?" The answer was quite typical... "Ah but it's the *way* we do it that's important." To an extent this is a reasonable reply but it strikes me that it is not a reasonable reply when it is the only one ever made to this sort of objection, no matter how many times you meet a problem by generating a new formalism, within the same computational paradigm, specifically to deal with it.

Formalism has very few limits on its possible coverage because its constraints are things like consistency and completeness.[3] We are left with a large amount of choice about what to modify when we come up against a problem. Given that our job is to simply design a formalism that covers a certain construction, we are almost guaranteed to be able to do it, and in many different ways. For example, if you want to have a very compositional approach but your representation of determiners is not well formed, adopt lambda abstraction; if you want to be able to represent the meaning of multi-sentence discourse but the variables in the different sentences are unrelated, simply invent some formalism for variable threading. Only technical problems prevent this and they are not restrictive enough to prevent you from augmenting the formalism in arbitrary ways guaranteed to cover the data. Intractability of the computational approach is met with a new tactic today. We spread out into more and more formalisms that each end up facing the same problems again and again. We are, in effect, reinventing a square wheel. The question of substance here is: "what is it about the computationalist program that allows this?"

The crisis manifests itself because of the nature of the traditional logical approach. The whole allure of formal systems in performing their original normative role of reforming and clarifying language is that they are very flexible. They are our conscious and carefully designed creations and we allow that we may augment and improve them as we see fit. This sort of arbitrarily adaptable tool is not the sort of thing very conducive to an open debate on the foundations of a subject. The reason being that you will never feel compelled to deny your basic principles when you have a tool that can always be modified in some way to cope with recalcitrant experience. The traditional approach is supported by the logical tools it has adopted and these tools have turned out to be fantastically methodologically elastic.

I think this is an instance of a more general observation. A normative or prescriptive system is one not totally constrained by the evidence, but one that seeks to constrain it. Thus such a system is designed to do violence to the way things are. Such systems are by nature reformative. As a result, they are designed to be very flexible and accommodating of the desiderata of a good prescriptive theory. The trouble is that the desiderata for a good prescriptive theory are not dependent on features of the evidence they intend to prescribe; that is the whole point of them. This makes them quite naturally unsuitable for a Cognitive Science having an *explicative* ingredient that marks it out from the purely descriptive engineering practices of Computer Science and to some extent, contemporary AI. The root of the formal, computational approach so apparent in semantics is exactly a prescriptive system. For example, if you have a theory X that encounters a problem piece of evidence Y and your theory is based on a prescriptively designed formalism, you will never have difficulty in bending the theory to fit the problem because the underlying formalism was designed with the independence of the features of the evidence and finished theory in mind. So the theory is not really constrained to features of the evidence. If you have a problem with the logical independence of the terms in supposedly analytic sentences like "All bachelors are unmarried," you have the tools to construct complex expressions to serve as meaning postulates or perhaps you might construct lexical decompositions to square the data with theory. All that is necessary to sanction this move is the formalism: the reason you perform an arbitrary formal operation to solve a problem is because you are able to do so.

Now, it may be objected that it is rather strange to suggest that a tool that can always account for the data is a *bad* thing. Well, in the face of such a tool, you can make roughly two responses. Firstly, admit that it is a positively *good* thing to be able to always account for the data. It means you are doing the right thing. This is of course might be a legitimate tactic although one we might worry about this along methodological lines akin to Popper's famous concerns regarding Freudian and Marxian theory for example. Unfalsifiable theories are methodologically suspicious. Secondly and more accurately in my opinion, if the tool was designed as prescriptive, reformative and idealized, being able to account for all the data is no longer a virtue. Its "success" follows from the nature of the tool rather from the connection a theory employing the tool has with empirical reality.

Once might be tempted to suggest that surely some evidence can legislate between differing camps in Cognitive Science? After all, most people hold that it is *empirical* after all. This consideration leads me to the second main source of divide within the Cognitive Science community: the nature of evidence.

3. The Evidential and Explanational Schisms

Formalist semanticists still take to be evidence, in the explanation of the semantic aspect of human cognition, the intuitions of native speakers or understanders. Thus, in this Chomskian vein, we take note of intuitions about quantification scope, anaphoric resolution, relative clause nesting etc. The more psychological and neuropsychological camps do not take this as evidence as such. Evidence there is reaction times, discrimination task performance and the like. This is quite serious as the differing camps can barely agree on evidence to disagree about. Part of the reason that the computational program constrains the field to such a degree is that is has a monopoly on what is to count as evidence. For example, I have heard it said many times that the neuroscientific data is all very interesting but is entirely the wrong sort of level of explanation we should be concerned with in Cognitive Science. It is "too low" a level of explanation to be scrutable. In particular, it is too low level to be input to traditional logical formalism. The evidential scruples of the computational program are, again, a product of its appropriation of the logical formalisms. Logical systems were designed to function as tools to standardize and disambiguate in the service of science. The perfect logical language would be clear, concise and paradox free. Logic undertook to reform the propositional expression of information within the sciences therefore concepts such as "proposition" and "deduction" are appropriate elements. Also, one of the central features of the natural sciences is that they are *written* disciplines. The primary mode of communication of scientific ideas is literary. As a result, there is a heavy emphasis on clarity and portability of expression. Such material should not be particularly contextually and indexically dependent. This contrasts starkly with natural spoken language which actively *depends* upon contextual and indexical information to an enormous degree. A formal ramification is that the "canonical language" is not well suited to the task of accounting for context. Workers in the field have typically attempted to account for context using the same formal tools, resulting in famous systems that have floundered due to computational explosion when attempts have been made to extend them beyond their toy domains. Commonly, the list of predicate/argument expressions that embody the formalist approach to "context" simply get too big too quickly. So, it seems clear that here is a case where the design constraints of the chosen formalism have become a burden for the computationalist paradigm. However, once you have adopted the formalism, you get its inbuilt desire for a particular type of explanation for free. This type of explanation is far from inclusive of all relevant levels one might like to explore and thus the space of possible research is forcibly and questionably restricted.

The standard account of the veracity of computationalist style explanations has to do with the necessity of certain types of constraint on patterns of behavior. Pylyshyn's famous argument is that if we have a level of explanation E whose explanatory elements are insufficient to constrain any resultant model of reality to reality, then we require another level of explanation E' to explain such constraints.[4] Thus, the computationalist program depends on the idea that the formal level of explanation is necessary to account for regularities in our semantic life. Pylyshyn passes over rather quickly the response that the constraints on our behavioral patterns are merely contingent and accidental, arguing that the symbolic level's exclusion of certain types of explanation is not exclusion *by definition*. But I simply cannot see how it could exclude other explanations in any other way since the formal tools are, in their inception, *prescriptive*: they are not designed with features of the world in mind so much as features of a formally attractive model of the world. Furthermore, there seems to me to be a principled way of achieving some of the more reasonable constraints that computationalists desire without necessitating the generation of more "levels of description."

It is well known but still often underrated feature of evolutionary theory that nature is economical. The reuse of existent structures is accepted as an enormously fruitful way of regarding the genesis of features of body and behavior. The functions of many biological structures are forced onto them given changes of environment and genetics. In keeping with this, one would expect and desire an explanation of the intricacies of language and semantics expressed in terms of preexistent structures and concepts. Contrastingly, the modern trend has been towards a *sui generis* treatment of this aspect of human capacities.[5] We are reasonably well informed about the kind of structures that give rise to motor behavior and have, over the past fifteen years, proposed models empirically supported by work in neuroscience.[6] Given the, hardly implausible, assumption that motor behavior precedes linguistic in the evolution of the human species, we might predict a reuse of structures and strategies found in the temporally prior behavior. This is exactly what neuroscience began to suggest over ten years ago.[7] However, the kind of representations posited as underlying the reuse of motor coordination constructs are emphatically not of the sort formalists in Cognitive Science are used to dealing with. Inter-methodological incommensurability results: a point addressed above. This is not the place to undertake a lengthy exposition of the technicalities of the more geometrical approach suggested by these considerations but I should point out two substantial elements in its favor.

Firstly, the formalisms employed are explicitly descriptive. Geometrical mathematics is designed to best fit data within well known constraints of consistency with other branches of mathematics. It is designed to model how things are and not how we might like them to be. Thus its basic concepts are nowhere near as pregnant with explanatory biases as those of formal logic; "point," "line" and "space" hardly press one in the direction of a particular view of mind and language as forcibly as "proposition," "predicate" and "object/meta language." Secondly, the computational and explanatory advantages of this method have historical precedent. Einstein's rendering of gravitation as a *structural* feature of spacetime as opposed to a force within a structure is the modern archetype of good explanatory methodology. Not only does this cohere with the post Duhemian concern with parsimony, it has very important pragmatic implications for Cognitive Science. Dreyfus and Winograd have long held that the computational explosion one encounters in attempting to model inference using symbolic representation is the result of fundamental ignorance of 19th century phenomenology. What is not currently appreciated is how far the geometrical approach goes towards dealing with this. The computational explosion is a result of the syntactic intractability of *semantic* relation. For example, the truistic character of "No bachelor is married" is not initially syntactically explicable as it depends upon the meaning of "bachelor." The standard way of dealing with this is to either allow a lexical decomposition of "bachelor" into "unmarried man" or to propose a meaning postulate meta-theoretically linking models containing "bachelor" and "unmarried man." The trouble is that this problem is ubiquitous in language and explicit delineations of such relations along either line are just not plausible if we are to consider the amount of processing that the brain must perform under such models. It is exactly the problem that Husserl and even Carnap faced in attempting to provide formal models of phenomena and fails for exactly the same reasons. However, in a geometrical model we have the opportunity of following Einstein in rendering semantic relationships as *structural* features of the space in which an representation might occur, thus obviating any computational penalties associated with explicit representation. There is not space to detail this approach here but a small example should suffice. If we have some eggs in a square box with the lid closed, we can say little about the relations between their positions: the space they inhabit is fairly orthogonal. However, put them in an egg-box and the situation is very different. We know automatically *from the structure of the space they now inhabit* certain things about them. For example, we know that none of them lie on their sides. We know that the distance relations between eggs obey transitivity as egg-boxes are made to be all alike (so they stack well) and this defines a metric on the "egg-box space." None of this has been derived by any explicit rules or inference: the "conclusions" are facilitated by the structure of the space the eggs now inhabit. Computationally speaking, we have relations for free in the same way that gravitational effects come for free in General Relativity. This is a way of caching out Dreyfus' insistence on our ability to "just see" certain semantic relations without the need for any actual *inference*.

This conception of Cognitive Science also addresses the linchpin of the formalist approach: the productivity of language and thought. Actually, productivity has never been a particularly strong argument for formalism as proponents of the view have admitted.[8] Infinite and even truly massive finite productivity is an idealization never actually realized. If this is the case, then we hardly need a recursive formalism: an iterative one will do as long as there are enough iterations to cover the life of any given human. Also, given that we link the need for a productive formalism to the productivity aspect of language and thought, we spend a lot of time having to invent restrictions on our formalisms to prevent them being *too* productive. I am thinking here of non-monotonic and strongly typed formalisms etc. No, a much stronger argument for formalism is the well known argument from the systematicity of language and thought. This is one of those constraints that Pylyshyn thinks necessitate a formal level of explanation as it is a constraint not reflected in, say, sub-symbolic explanation. I think this is no longer the case. We are now beginning to see how the complexities of the structure of spaces employed in so-called "geometrical" models may well be able to provide a non-symbolic circumscription that fits in with the largely justified systematicity desiderata of the symbolist.[9]

Not only are there differences in what counts as evidence in the field: there is also a rather stark difference in what is to count as an explanation. This is brought out clearly in the disagreements between symbolists and connectionists. The latter are often happy to allow parts of their networks to have no interpretation under a particular theory.[10] The symbolists, however demand that their evidence be systematically interpretable and even go to great lengths, as in the case of providing "meanings" for determiners by using lambda abstraction, to ensure that it is. This desire for explanations and accounts that have something to say about every stage and aspect of a process seems to me to be a clear consequence of one of the central theses of semantics; that of compositionality. Drawn from Frege (indeed sometimes called "Frege's Principle") this requires that the meaning of a complex expression is a function solely of the meanings of the more basic expressions that comprise it. As a consequence, those accepting this principle prefer explanations of *X* that are parasitic upon explanations of parts of *X*. Thus every aspect of an account must be made clear in terms of the theory. Again, this seems to be a consequence of features of the adopted formalisms. The desire for compositionality is originally a formal one. We shall force this on our data; indeed Frege explicitly gave referents to non-referring expressions to accomplish this. This is not because non-referring expressions really do have referents and our formalism is telling us this, but because we would like them to have, in the name of formal consistency. This is a clear case of reformative formalism. As a result, I do not think it a coincidence that formalists came to believe in the compositional nature of mental representation. Their formal tools had this built in when they were adopted.

A general but underappreciated feature of formal sciences is the way in which their formalisms generate paradigms of explanation that mirror formal features. Formalisms have the reputation of being tools by which one implements a theory. As a result, their features as naturally seen as posterior to theory: a result of theory. Often, this inchoate view is mistaken. Adoption of a formalism promotes a two way process in which assimilation of theoretical distinctions into the formalism is only a proper half. The other half is the assimilation of formal distinctions into the theory. The traditional Fodorian modularity theses regarding the elements of inference are,[11] I think, as much a result of the formal distinction between object and meta-language as any putatively empirical evidence. Problematically, the formalism is not guaranteed to be a provider of good cues. Indeed, in Cognitive Science, given the attitudes of the formal progenitors, quite the opposite is seen to be the case. The maxim is: when using your formalism, be careful it does not use you.

Now, it is not always a bad thing to have your formalism suggest novel and unlooked for features of your data: indeed it is often a striking methodological bonus when this happens. However, the case before us is unduly troublesome in at least the following two respects. As mentioned above, the dictates of a formalism decidedly hostile in its inception are somewhat more suspect than a formalism designed to do what you are using it for. A formalism that suggests one finds theoretical significance in, for example, an object/meta language distinction had better be sympathetic to your overall purpose.[12] Otherwise, you are at the very least straining the application of the formalism to its natural limits. Secondly, the data we are considering here is of a quite different sort to that which we find in the natural sciences. The problem for the formalist in general, is that what is taken to be evidence shares a perniciously symbiotic relationship with the theories it is meant to inform. This is not to suggest that one should hope for a reinstatement of the long forsaken observation/theory language distinction. Rather that because native intuitions are such a strange sort of evidence, they are particularly prone to self-fulfilling prophecy effects. When you have tried to decide a few times whether or not a sentence has ten or twelve scopal readings, your intuitions become so very confused that they simply do not have the basic feature required of evidence: they are not particularly stable. If you want to test the current water level the last thing you want to stick your yardstick in is a raging sea. The battle between a reformative formalism and unstable data is a foregone conclusion. You always succeed in covering the data. By giving oneself by definition, in the manner of Chomsky, a theoretically stable but empirically inaccessible level of "competence," one does nothing to allay fears that the formalism is excessively driving the theory. Rather, the fears are confirmed.

So, the problem is compounded: not only do we have a formal assumption that allows us to continue along the same road in the face of every adversity; we also have a notion of evidence that is unstable enough to support serious critique of any particular theory. This adds up to a seemingly principled façade resulting in a difficulty in entertaining fundamentally different approaches: the position that Cognitive Science and particularly linguistic semantics finds itself in today.

Notes

1. Frege (1882).
2. Tarksi (1931), in Tarski, (1956).
3. In higher order systems, we often have to do without even these.
4. Pylyshyn (1984: 35–38).
5. A recent example of this is McDowell (1994).
6. Churchland & Sejnowski (1992) provides a good overview of recent work.
7. See, for example, Pellionisz & Llinás (1982).
8. See, for example, Fodor (1987).
9. See Gärdefors (1990) for the beginning of such an approach.
10. There is a current trend towards providing a type of compositional treatment for connectionist models nowadays. See, for example, Gärdenfors (1993). This strikes me as a consequence of the monopoly that the symbolic paradigm has on the concept of explanation.
11. See, for example, Fodor *et al.* (1980).
12. As is the case in Fodor *et al.* (1975); Fodor *et al.* (1980).

References

Churchland, Patricia S. and Sejnowski, Terrence J. 1992. *The Computational Brain.* Cambridge, MA: MIT Press.

Dreyfus, Hubert L. 1992. *What Computers Still Can't Do*. Revised edition of *What Computers Can't Do,* 1972. Cambridge, MA: MIT Press.

Fodor, Jerry A. 1987. Why there still has to be a language of thought. In *Psychosemantics*, 135–167. Cambridge, MA: MIT Press.

Fodor, J.D., Fodor, J.A. and Garrett, M.F. 1975. The psychological unreality of semantic representations. *Linguistic Enquiry* VI(4), 515–531.

Fodor, J.A., Garrett, M.F., Walker, E.C.T. and Parkes, C.H. 1980. Against definitions. *Cognition* 8, 263–367.

Frege, G. 1882. Uber den wissenschaftliche Berechtigung einer Begriffsschrift (on the scientific justification of a conceptual notation). *Zeitschrift für Philosophie und philosophische Kritik* 81. English translation by Bartlett in *Mind* 73, 1964.

Gärdenfors, Peter. 1990. Induction, conceptual spaces and ai. *Philosophy of Science* 57, 78–95.

Gärdenfors, Peter. 1993. How logic emerges from the dynamics of information. Lund internal paper.

McDowell, John. 1994. *Mind and World.* Harvard University Press.

Pellionisz, Andras and Llinás, Rodolfo. 1982. Space-time representation in the brain. The cerebellum as a predictive space-time metric tensor. *Neuroscience* 7, 2949–2970.

Pylyshyn, Zenon W. 1984. *Computation and Cognition*. Cambridge, MA: MIT Press.

Tarski, Alfred. 1931. The concept of truth in formalised languages. In Tarski, 1956.

Tarski, Alfred. 1956. *Logic, Semantics and Metamathematics*. Translation by Woodger. Oxford University Press.

Biomolecular Cognitive Science

Ajit Narayanan
Department of Computer Science
University of Exeter

1. Background

Cognitive Science, quite simply, attempts to provide solutions to the question of how mind and brain are related or, more generally, what constitutes mind/brain. *Classical* cognitive science (CCS), together with its subdiscipline of artificial intelligence (AI), is based on Newell and Simon's (1976) explicit commitment to the "Physical Symbol System Hypothesis" — the idea that all intelligent action and behavior can be necessarily and sufficiently described and explained by symbols and rules operating on those symbols, where the rules themselves can have symbolic form. Furthermore, these rules and symbols must be realized in any system for which claims of intelligent action and behavior are made.[1] AI's concern has been with *computational representations* of physical symbol systems.

CCS and AI have been attacked on the following three grounds: (a) that rule-following by itself is not sufficient (and may not even be necessary) for intelligence, awareness and consciousness; (b) that because CCS and AI are anti-materialist and perhaps anti-reductionist in nature they cannot explain how brain gives rise to mind and therefore cannot provide adequate accounts of mind/brain; and (c) that CCS and AI, because they succumb to the same formal limits that apply to computation and algorithms, cannot account for certain types of mental processes which fall outside the class of what can be computed.[2]

With regard to (a), the strongest expression of this objection to CCS and AI has come from Searle and his Chinese Room Argument (Searle 1980). The best reply to the actual Chinese Room scenario is the Korean Professor Argument (Rapaport 1988) which identifies a weakness in the Chinese Room scenario (the person in the room understands the instructions to be followed) before re-describing the scenario in a form acceptable to CCS and AI.[3] More generally, though, AI addresses this problem through a form of "Systems reply": intelligence, awareness and consciousness arise from computational processes and interactions between these processes. Whether these processes are mental, physical or behavioral is irrelevant, in that the system as a whole moves through various states, where the next state of the system is determined by the current state of the system and any input it receives. The stress here is on functionality and cognitive architecture (Putnam 1967; Fodor and Pylyshyn 1988): The states a system goes through are *representational* states, and a cognitive architecture is an architecture of representational states which involves the precise nature of the representations and the operations performed over them.

With regard to (b), neuroscientists claim that an understanding of the brain is required for any account of mind, where the claim is supported by evidence that so far it has not been possible to find an entity with a mind which/who does not also have a brain. In response to connectionist attacks (Rumelhart and McClelland 1985; Rumelhart, McClelland *et al.* 1986) Fodor and Pylyshyn (1988) further refined CCS to identify three characteristics which representations in any proposed cognitive architecture have to satisfy: systematicity (the ability of a system to produce/understand some expressions is intrinsically connected to the ability to produce/understand certain others), productivity (the ability to produce/understand expressions not previously encountered), and compositionality (the ability of an item to make the same semantic contribution to each expression in which it occurs).

There have been a variety of attempts to provide connectionist representational architectures which satisfy these characteristics (e.g. van Gelder 1990; Bodén and Narayanan 1993; Niklasson and Sharkey 1994; Christiansen and Chater 1994; Niklasson and van Gelder 1994).[4] But there are two types of neuroscience. On the one hand, *reductionists* in general accept that, even after reduction to a neuroscientific basis, mental processes do exist and can be described in their own terms.[5] This is to be contrasted with eliminative neuroscientists, who believe that the sort of reductionism canvassed by reductionists does not go far enough: "Eliminative materialism is the thesis that our common sense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience" (Churchland 1981: 206). This leaves neuroscientists with the problem of how to account for mind. The concept of emergentism is often appealed to at this point: a collection of relatively simple neuronal units, communicating with neuronal units at neighboring levels, together perform a global (holistic) computation that none of the individual units, or linear combinations of them, could do alone. Emergentism is the idea that what are called higher level cognitive processes can be accounted for by their emergence from the neurocomputational substrate. However, there is as yet no clear neuroscientific account of emergentism, except for references by way of analogy to the way that the microstructure of a physical object (in terms of atoms, molecules and lattice structure) can give rise to macro level physical properties (e.g. hardness).

The stance of CCS and AI on these issues has been that it makes as much sense to ask for details of the way the brain works when trying to understand the mind as it does to ask for details of hardware when trying to understand how a program works. That doesn't mean that an *implementation* of an algorithm or mind does not require hardware or a brain, respectively; rather, what is claimed is that details of the hardware/brain do not add anything to our algorithmic/mental accounts. The core question for neuroscientists, "How can the brain as material object evoke consciousness/mind?" can only be answered by appealing to representational states, and CCS and AI are best placed to offer an account of representational states, goes the argument.

With regard to (c), the Mathematical Objection (MO) is that machines will never be able to do everything human minds can do (Lucas 1961). This is because Gödel showed that any formal system of a sufficiently powerful kind cannot be both consistent and complete at the same time. This means that there will always be one statement which, if true, cannot be proved, and if proved, cannot be true. Since a computer and its program are an instantiation of a formal system, it follows that for any AI computer there will always be one statement (called the Gödel Formula) which the computer cannot see as true (or provable) but which we humans can see is true (or provable). Proposers of the MO claim that this argument prove that machines can never do everything that humans can do, that machines will always be one step behind human reasoning.

AI has traditionally replied to the MO in a variety of ways. For instance, the criticism that a computer cannot "jump out of the system" assumes that systems are logically separated onto separate levels, with simple systems at the bottom and increasingly complex systems at higher levels. But in CCS and AI the brain is at the bottom level, and the brain, if it is describable mathematically at all, will have a complex mathematical description. The only way to understand the brain is to "chunk" it on higher and higher levels, thereby *losing precision* until perhaps at the higher levels we have "informal" systems (Hofstadter 1979). That is, levels in a mathematical proof and levels in AI are not the same. Therefore, what a mathematician and an AI researcher jump out of and into are different also.[6]

2. Biomolecular Foundations of Mind/Brain

However,a fourth objection is now surfacing which presents serious difficulties for both CCS and neuroscience. The fourth objection — that consciousness is *biomolecular* and that any account of the mind/brain which does not take into account the behavior of biomolecules is doomed to failure — undercuts neuroscience which stresses *neurons* as the primitive computational element as well as classical cognitive science which stresses a cognitive system passing through various representational states. As Penrose (1994: 357) says:

> If we are to believe that neurons are the only things which control the sophisticated actions of animals, then the humble paramecium [a single-cell, eukaryotic organism belonging to the kingdom *protista*] presents us with a profound problem. For she swims about her pond with her numerous tiny legs ... darting in the direction of bacterial food which she senses using a variety of mechanisms, or retreating at the prospect of danger, ready to swim off in another direction ... Moreover, she can apparently even *learn* from her past experiences ... How is all this achieved by an animal without a single neuron or synapse?

The implication, quite simply, is that neuroscientists have got it wrong if they claim that networks of neurons (single cells) adequately account for mind/brain: while a cell is the basic unit of living systems, this does not mean that the cell is primitive.

Molecular computing, which stands in the same relationship to biomolecular science as AI does to CCS and connectionism to neuroscience, is the computational paradigm derived from and/or inspired by biomolecular processes within cells (Carter 1984; Conrad and Liberman 1982; Hameroff 1987).[7] So, what is molecular computing, and can it and its parent science provide an adequate account of mind/brain? A brief description of cell structure and function is required at this point.

A cell — typically 10–30 millionths of a meter across for humans — contains many specialized structures called organelles. The relevant ones here are the cell membrane (controls passage of substances into and out of the cell and encloses cell organelles as well as cell substances), cytoplasm (serves as a

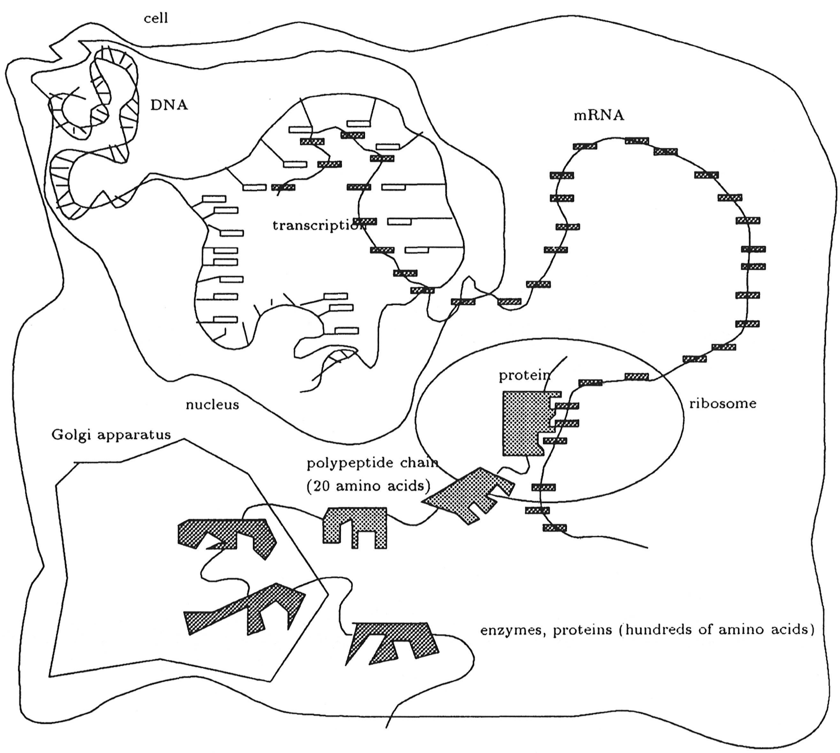

Figure 1. Only the relevant parts of the cell are shown here: the nucleus which contains DNA and RNA, ribosomes where protein construction (translation) takes place using the 20 basic amino acids, and the Golgi apparatus where individual amino acids are modified slightly to produce the variety of amino acids essential for life. A body cell is typically 10–30 millionths of a meter long and wide, and it is estimated that we have several trillion of such cells (for skin, muscles, liver, blood, heart, brain (a neuron is a brain cell), etc.). Each such body cell contains the full set of 46 chromosomes (discrete molecular structures of DNA) inherited from our mother and father (23 in each case, via sex cells). The "straight-line" length of the DNA in one cell is estimated to be 2 meters, which demonstrates the tightly packed nature of the chromosomes and their thinness. It is also estimated that the 46 chromosomes code for between 75,000 to 100,000 genes for humans, using about 8 billion bases (nucleotides). On average, about 100,000 bases are required for coding a gene, although this figure varies greatly from a few hundred to a few hundred thousand.

fluid container for cell organelles and other cell substances as well as assists in the transport of substances within the cell), nucleus (directs all cell activity and carries hereditary information), endoplasmic reticulum (serves as a transport network and storage area for substances within the cell), ribosome (manufactures different kinds of cell protein), Golgi apparatus (packages protein for storage or transport out of the cell), lysosome (digests or breaks down food materials into simpler parts and removes waste materials from the cell), mitochondria (serve as the power supply of the cell by producing ATP — adenosine triphosphate — which is the source of energy for all cell activities), microtubules (serve as the support system or skeleton of the cell) and microfilaments (assist in cell motility). Each organelle performs one or more special tasks to keep the cell alive. All the information directing every cell function is stored in large DNA molecules found in the nucleus.

A cell cannot function without DNA. The information it contains must be made available somehow to the rest of the cell as well as be passed on to all new cells. Although each cell contains the full complement of DNA, through some process which is not yet clearly understood certain parts of the DNA are switched on or off within cells, resulting in different types of cell producing different proteins for normal growth and functioning of the organism as a whole. The process by which the information in the DNA is carried out to the rest of the cell is through messenger RNA strands which leave the nucleus and attach themselves to the ribosomes, which then produce the protein for export from the cell (Figure 1). What is remarkable is that the DNA are large molecules made up of combinations of only four types of nucleotides — adenine, guanine, thymine and cytosine (called A, G, T, and C, respectively). It is estimated that the DNA in each one of our cells contains about 8 billion nucleotides, spread across 46 chromosomes (discrete molecular structures of DNA), each one of which takes the shape of a double helix. If all the DNA in one cell were stretched end to end, the length is estimated to be about two meters. Messenger RNA bang into these chromosomes and unzip part of the molecule, make a *complementary* copy of a certain length of the molecule, before leaving the nucleus for the ribosomes and protein manufacture. The process of DNA being mapped into mRNA is called *transcription*, whereas the process of duplicating all chromosomes is called *replication.*[8] Ribosomes produce the appropriate amino acids from the mRNA. For instance, the mRNA triplet *GCU* (guanine — cytosine — uracil), which is an mRNA transcription of the DNA triplet *CGT* (cytosine — guanine — thymine), is mapped onto the amino acid *alanine* by ribosomes.

It may appear from the above that the transfer of information from the nucleus to the rest of the cell is a highly organized affair. This is not correct. Random collisions millions of times a second between RNA polymerase (an enzyme, which is a large protein which helps make and break bonds) and the

DNA eventually lead to the RNA polymerase running into certain sequences of bases and latching onto them. These sequences of DNA bases are recognized by the RNA polymerase as start positions for transcriptions. The RNA polymerase then unravels the appropriate part of the DNA double helix. Free-floating bases in the nucleus attach themselves to the revealed DNA bases, forming a sequence which becomes the messenger RNA. The double helix is re-formed as transcription continues along the unravelled DNA molecule. When a terminating sequence of bases is found in the DNA, the resulting messenger RNA is dispatched to the ribosomes, where combinations of three bases at a time in the messenger RNA are used to produce one of 20 different amino acids. Sequences of these amino acids (varying in length from a few hundred to a few thousand) are called polypeptide chains, which are packaged in the Golgi apparatus and then secreted from the cell for use by other cells in the organism. These polypeptide chains therefore "represent" the sequence of bases unravelled in the DNA molecule (Figure 1). Again, it may appear that the production of polypeptide chains out of individual amino acids is a highly organized affair. This again is not true: there are so many millions of molecular collisions each second within ribosomes during polypeptide production that some of these must be the correct ones for the proper production of the polypeptides.[9] For instance, appropriate polypeptides (proteins/enzymes) for continually producing hair of a certain color for an individual are transferred from the individual's DNA in certain specialized hair-production cells.

3. Implications for Cognition

So, what happens to the enzymes/proteins produced by ribosomes and the Golgi apparatus? Proteins (enzymes) carry out many vital functions in living organisms. As structural molecules, they provide much of the cytoskeletal framework of cells. As enzymes they act as biological catalysts that speed up the rate of cellular reactions. The chemical composition of one of our cells could be placed in a test-tube and observed, We may, after some time, notice some chemical reactions naturally occurring in the test-tube. There will be a long delay because the activation energy required to start a chemical reaction acts like an energy barrier over which the molecules must be raised for a reaction to take place. An enzyme effectively lowers the activation energy required for a reaction to proceed. An enzyme locks onto a molecule, starts a reaction, and then is released unchanged. The rate of enzyme combination and release is called the *turnover rate* and is about 1000 times a second for most enzymes, with variation between 100 per second and 10 million per second. The increase in reaction rate achieved by enzymes ranges from a minimum of about a million to as much as a trillion times faster than an uncatalysed reaction at equivalent concentrations and temperatures. From this it can be seen that the process of enzyme/protein production, as determined by our DNA, is absolutely critical to our continued well-being, otherwise we as chemical beings would not produce chemical reactions fast enough to keep us alive (e.g. respiration, digestion).

What inheritance now means, according to biomolecular science, is the set of genes (DNA) which code for the production of appropriate enzymes which increase the rate of chemical reactions in our cells, where the nature and rate of reactions is determined by the nature of the enzymes. We are all essentially the same chemically: what differs is the enzymes produced by the DNA inherited by our parents and other factors (e.g. mutation of individual bases and genes by random means), and these enzymes control cellular processes differently for different people, thereby leading to different physical characteristics.

The applications of molecular computing are quite clear in the area of biomedical research. For instance, cloning is the process in which a diploid cell divides and produces a whole new organism through complete DNA replication (rather than the nuclei of two haploid cells merging to produce off-spring).[10] Also, various inborn errors of metabolism and chromosome errors which give rise to genetic diseases can be explained as errors in transcription or replication and through DNA mutation, such as sickle cell anemia (reduction in the solubility of hemoglobin in the blood), Tay-Sachs disease (absence of specific enzymes that hydrolyse specific lipid molecules), diabetes mellitus (insulin deficiency), hemophilia (improper clotting of blood), phenylketonuria (associated with mental retardation) and albinism (the production of skin pigment *melanin* is blocked). Viral infections (colds, flu, measles, chickenpox and mumps, for example) can be explained at a deep level and resulting computational models can generate hypotheses concerning their evolution and treatment.[11] Cancers of various sorts exhibit a wild, uncontrolled growth, dividing and piling over each other in a disorderly arrangement and pushing aside the normal cells in a tissue. Similarly, computational models can provide a useful service here in prediction and treatment. But what are the implications of molecular computing and its parent, biomolecular science, for mind/brain? Although a cell has a functional architecture and performs many functions, each of which is determined by the DNA information within a nucleus, the idea of a cell moving through various representational states as it processes information, where there representational states (as required by CCS) involve the manipulation of symbols, does not sit easily with the facts. Molecular computing is essentially a copying (translation and replication) process. However, if molecular computing and biomolecular science are to offer alternative accounts to CCS and neuroscience, there must be some method by which elements of molecular computing are tied up with representations, information processing and consciousness.

There are a variety of proposals in the molecular computing literature concerning the way that cells could give rise to information processing and consciousness. They can be split roughly into two types: the first type deals with the way that any cell can be regarded as an information processing device, and the second deals specifically with brain cells and attempts to use properties of neurons for accounting for consciousness. Among the proposals of the first type are claims that cells represent information through (i) reaction diffusion systems (chemical reaction waves within a cell propagate at uniform speed and interact with other waves within the cell to produce complex patterns (Conrad and Liberman 1982; Winfree and Strogatz 1984)) (ii) cellular automata (a large number of identical cells connected in a uniform pattern and communicating only with other cells in their neighborhood operate collectively to produce complex behavior (von Neumann 1966)) and (iii) the protoplasm (dynamic activities of cytoskeletal structures including cytoplasmic microtubules within a cell produce rudimentary consciousness (Hameroff 1987; Penrose 1994)). Among proposals of the second type are (i) holograms (the brain perceives sensory information by analysis of the interference of neural firing frequencies, resulting in a domain in which space and time are enfolded (Pribram 1986)), and (ii) cytoskeletal activity, but this time within neurons and at a quantum mechanics level (Hameroff 1987; Penrose 1994).[12] However, it must be said that all these proposals are highly speculative, leaving CCS with its stress on computation and neuroscience with its use of mathematically rigorous connectionist networks in the lead as far as clear proposals are concerned.

4. The "Mind Gene"

The question now is: Is there a mind gene? That is, is there a part of our chromosomes which produces enzymes/proteins which, when released in, say, neurons, give rise to consciousness and mind? The current approach to this question consists of appealing to an *evolutionary* account. As Crick (1994: 12) says about language:

> ... the understanding of the evolution of language will not come only from what linguists are doing, but from finding how language develops in the

> brain ... and then finding the genes for it and trying to work out when those genes came in evolution.

Two approaches to this question can be predicted. The first depends on a contextual approach where, for instance, to account for, say, a type of sensation is to identify which part of one's DNA (hereditary information) is responsible for producing the polypeptide chains associated with that sensation and then to derive an evolutionary account based on neighboring DNA code. An account of desire may be based on identifying which part of one's chromosomes is responsible for producing the chemical proteins/enzymes associated with desire and then determining what is on either side of the DNA for desire. It may be that on one side in the chromosome is the code for producing polypeptides associated with goal-motivated behavior, and on the other the code for producing polypeptides associated with plan-producing behavior. An evolutionary account of desire would then be based on some story which related goals to plans by means of desire: At some stage in the evolution of consciousness, it was found beneficial for organisms to have desires as a way of bridging the gap between goals and plans for achieving those goals, for example. Such contextual answers are subject to the criticism that whole genes may be moved from one part of a chromosome to another through random displacement or peculiarities of DNA folding.

The second type of answer depends on identifying homologous (common ancestor) DNA. The gene for desire may be found to contain significant amounts of DNA associated with, say, goal-motivated behavior as well as its own specialized DNA. Desire can then be explained as having evolved from (inherited) goal-motivated behavior but also to have specialized with respect to goal-motivated behavior.[13] Such homologous answers are subject to the criticism that specialized genes may contain significant exceptions to what their common ancestor gene contains and may also inherit from more than one common ancestor.[14]

Irrespective of the approach adopted, a biomolecular approach to cognitive science implies a different way of looking at mind. A mind genotype is the genes a person has for mind, whereas a mind phenotype is the expression of these genes in actual thought. One possibility here is to predict actual thought processes of a person from a knowledge of their mind genotype, should such a genotype be discovered in our genes. However, mind phenotype may not be easy to predict, even with knowledge of mind genotype. While some phenotypes are discrete and can be predicted from their genotypes (e.g. blood type), other phenotypes are continuous (e.g. height) and may be dependent on environmental factors (e.g. nutrition). Similarly, it may be argued that mind phenotype is not like blood type but more like height.

Also, just as genotypes are inherited from parents who in turn inherited from their parents, and so on, mind genotypes, if they exist, must be inherited from parents and their parents … The genetic make-up of one's mind consists of bits and pieces of mind genes of many ancestors, in the same way that color of hair or eyes consists of bits and pieces of inherited ancestor genes.

This raises the question of how many different types of mind genotype there are, and how they are manifest in mind phenotype. It is possible that a certain thought I have is of the same phenotype as a thought one of my ancestors (e.g. my grandmother) had, from whom I have inherited some bits of mind gene. But what distinguishes my thoughts from my grandmother's is my genotype, which is inherited from a wider pool of genes than just my grandmother. However, from a biomolecular point of view, the way my mind genotype is expressed in actual thoughts may be similar to one of my ancestors. Taking this to its logical conclusion, it could be argued that every one of my thoughts is an expression of certain aspects of my mind gene which in turn has been inherited from ancestors. I really do think like my grandmother, for some of the time anyway. Then I think like my mother, for some of the time, and so on. Why only some of the time? That's because during those times that part of my mind gene which I have inherited from my grandmother or mother is being transcribed to produce enzymes which, when active in my neurons, produces certain thoughts which are different in type from thoughts produced by enzymes transcribed from parts of my mind gene which I've inherited from, say, my father.

It's important to stress the difference between genotype and phenotype, and the continuous range of phenotype expressions possible for discrete genotype values. While mind is determined by genes, the expression of those genes in actual thought will be dependent on other factors also, such as current and previous experience, nutrition, and so on. The issue here is therefore not full genetic determinism. Rather, what is at issue is the classification of mind into distinct types, as given by the mind gene(s).

Of the estimated 75,000 to 100,000 genes in the human genome, many thousands have been identified and located on specific chromosomes. So far, there has been no identification of a mind gene, i.e. a specific location or set of locations in and across chromosomes which code for mind, but then it is not clear that molecular biologists are looking for a mind gene as such or would recognize it. A vast proportion of our DNA does not code for specific enzymes and therefore can be regarded as non-genetic. Their purpose is regarded as nonfunctional (except that such redundant DNA can be used for identifying individuals via DNA fingerprinting: the pattern of repeating "redundant" DNA in you is different from in me).

5. Conclusion

What the above has shown is that Rumelhart and McClelland (1985) were surely right when they state: "[T]here's more twixt the computational and the implementational than is dreamt of, even in Marr's philosophy ..." (1985: 196). The problem is that each level, to lend credence to its claim for accounting for mind/brain, posits a form of computation and representation not just appropriate for that level but also necessary and perhaps sufficient for explaining cognitive phenomena at that level. CCS proposes symbolic computation, algorithm and representation, neuroscience proposes mathematically based extraction of information and knowledge contained in connectionist networks (by means of hyperplane analysis, for example), and biomolecular science proposes molecular computing which is based on biomolecular processes within the cell nucleus. The recent interest in mind/brain issues shown by leading figures in the physical and biomolecular sciences indicates that cognitive science and neuroscience no longer have the field to themselves. There is already competition between the biomolecular and physical sciences as to which is going to prove to be the more suited for accounting for mind/brain.

The most immediate implication for CCS is that the notion of computation, tied as it is to the concepts of rule-following, algorithm, effective procedure and TM-computability (areas attacked by objections [a] and [c] described earlier), goes back into the melting pot. What may emerge is a concept of computation which is tied more closely to biomolecular principles (transcription, translation, replication, mutation, and so on) than to a formally specifiable and repeatable sequence of steps to reliably achieve a task. With regard to neuroscience (objection [b]) the most immediate implication is that neurons are at too high a level, and so any neuroscientific account of mind/brain in terms of layered networks will not be accurate. What is needed is a clearer understanding of the internal workings of neurons in biomolecular terms. Physical scientists may argue that biomolecular computing is still at too high a level and that its own computational arm, quantum computing (e.g. Deutsch 1992; Menneer and Narayanan 1995) provides a more appropriate computational level. A radical physical scientist may also claim that consciousness/thought is a feature of the brain's physical actions where these physical actions cannot even be adequately expressed in any computational terms (Penrose 1994) — the physical theory has no computational arm. The scientific foundations of mind/brain are currently up for grabs.

But the anti-materialist and perhaps anti-reductionist nature of CCS (objection [b]) may lead to it becoming isolated because of its unwillingness to accept computational paradigms and representations of levels lower than the algorithmic as real alternatives for an account of mind/brain. What is being proposed in this paper is that CCS should adopt, for the purposes of scientific hypothesizing, biomolecular paradigms. Even if this level is proved ultimately to be wrong, at least CCS will be contributing solutions to a fresh range of problems, some of which (e.g. explaining the biomolecular basis of diseases) have profound implications for humanity. More interestingly, though, if CCS adopts biomolecular paradigms, they can cut the ground away from under connectionists' feet by pointing out that, while networks of neurons may well perform certain tasks non-symbolically, within each neuron there are processes which can be described symbolically. Such processes are described using the symbol structures and processes of biochemistry (e.g. nucleotides, transcription, replication, enzyme production) and physical chemistry (e.g. molecule construction out of atoms, molecule folding), even if it is not currently clear whether these symbol structures and processes are computational. Nevertheless, it can pointed out that nonsymbolic behavior at the level of networks rests fundamentally on biochemical symbol structures and processes within each neuron making up the network. Nonsymbolic processes at the neural network level could be emergent properties of symbolic processes at the individual neuron level — an interesting twist.

Notes

1. Rules can be implicit rather than explicit, but symbols must be explicit (Fodor and Pylyshyn 1988).
2. Turing (1950) identified early versions of these objections as "The Argument from Informality of Behavior" and "Lady Lovelace's Objection" (for [a]), "Argument from Continuity of the Nervous System" (for [b]), and "The Mathematical Objection"(for [c]).
3. Imagine a professor in Korea who understands no English and who relies on the best available translations of Shakespeare's work in order to write, in Korean, deep, penetrating analyses of Shakespeare's plays. The Korean professor's writings are translated into English by translators and published in the best Shakespearian journals, leading to world recognition of the quality of the Korean professor's writings. If you

can imagine this, then, goes the argument, this is what CCS and AI are all about: the ability to reason and be creative, within a language that can be understood, even if that language needs translating for other to understand.

4. Searle (1987) proposed a form of biological materialism, where variable rates of neuron firing relative to different neuronal circuits produce all the different types of mental life we humans experience: "… mental phenomena, whether conscious or unconscious, whether visual or auditory, pains, tickles, itches, thoughts, and the rest of our mental life, are caused by processes going on in the brain" (p. 220 — stress removed), where these processes are "real biological phenomena" (p. 217).
5. This is analogous to a high-level program, even after being compiled into machine code, still existing as an entity in its own right as a textual entity about which certain judgements can be made (e.g. its complexity, structure, design).
6. Recently, another dimension has been added to the debate by Penrose (1989, 1994). CCS and AI have got it wrong, he argues, when associating thought, reasoning, consciousness and awareness with algorithms. Rather, algorithms, if they have any part to play at all in human intelligence, play a part at the "unconscious" or "subconscious" level, perhaps as a result of thousands of years of evolution. It is the role of consciousness to be non-algorithmic, i.e. to apply common sense, judgement, understanding and artistic appraisal to the results of our (unconscious) algorithms.
7. Hofstadter (1979) seems to have predicted this growing interest in biomolecular processes when he presented a simplified molecular computing system called Typogenetics.
8. Transcription differs from replication in that transcription involves the use of a fifth base — uracil — which is mapped onto by thymine. That is, during transcription, instead of DNA thymine being mapped onto mRNA adenine, DNA thymine is mapped onto mRNA uracil.
9. What has been described applies only to diploid cells, which are all body cells apart from the haploid, or sex, cells. In diploid cells, the nucleus of each body cell contains the full set of hereditary information. When a new diploid cell is formed, complete copies of all the chromosomes must be made through a process known as DNA replication.
10. Both translation of replication involve the unraveling of the DNA strands and the use of one strand as a template for joining together nucleic acid units in the proper sequence. After translation, the DNA recoils to its original form, but after replication the two original strands have separated, each with a new complementary partner strand.
11. A virus is a set of genes packaged inside a protein coat. Inside the coat of most human viruses is a single strand of DNA, coded to reproduce itself. A virus borrows the ribosomes of cells in a host organism to make proteins. New virus particles are formed which break out of the cell to infect other cells. A cell infected with a latent virus (unusual for humans) shows no sign of infection for a long time and may even have made many copies of itself (with the latent virus implanted in its DNA) before something triggers the virus into activity. AIDS is a highly unusual (for humans), latent retrovirus which injects RNA (and not DNA) into T4 white blood cells (an important

cell in a human's immune system). Reverse transcriptase then converts the RNA into DNA which is subsequently inserted into the T4 cell's normal DNA. White blood cells divide normally, carrying their infected DNA into new copies of themselves. When the latent infection is triggered, the emission of new copies of the AIDS virus causes the T4 cells to die, leading to a severe breakdown of the immunity system.

12. Interestingly, Hameroff and Penrose have opposite views on the adequacy of quantum computing (Deutsch 1992) for accounting for quantum mechanical effects, with Hameroff seeing no reason in principle why quantum computing should not provide adequate computational accounts of quantum behavior and coherence within microtubules, and Penrose believing that quantum computing, because it is still essentially computational, cannot in principle adequately account for how microtubules give rise to consciousness.
13. Object-oriented computational paradigms (e.g. the use of C++, CLOS, ONTOS), where inheritance is the basic mechanism for sharing information, can be predicted to be useful for modelling evolutionary accounts.
14. See Al-Asady and Narayanan (1993) for an overview of the problems associated with multiple inheritance with exceptions.

References

Al-Asady, R. and Narayanan, A. 1993. More notes on 'A clash of intuitions.' *Proceedings of the 13th International Joint Conference on Artificial Intelligence* (IJCAI–93), 682–687.

Bodén, M. and Narayanan, A. 1993. A representational architecture for nonmonotonic inheritance structures. In *Proceedings of the International Conference on Artificial Neural Networks* (ICANN–93), S. Gielen and B. Kappen (eds), 343–349. Amsterdam: Springer Verlag.

Brown, J. 1994. A quantum revolution for computing. *New Scientist* 1994, 1–24.

Carter, F.L. 1984. The molecular device computer: Point of departure for large scale cellular automata. *Physica* 10D, 175–194.

Christiansen, M.H. and Chater, N. 1994. Generalization and connectionist language learning. *Mind and Language* 9(3), 273–287.

Churchland, P.M. 1981. Eliminative materialism and propositional attitude. *The Journal of Philosophy* 78, 67–90. Reprinted in W.G. Lycan (ed), *Mind and Cognition: A Reader*. Oxford: Blackwell. The page reference is to the reprinted version.

Conrad, M. and Liberman, E.A. 1982. Molecular computing as a link between biological and physical theory. *Journal of Theoretical Biology* 98, 239–252.

Crick, F. 1994. Interview with Jane Clark. *Journal of Consciousness Studies* 1(1), 10–24.

Deutsch, D. 1992. Quantum computation. *Physics World* 5, 57–61.

Fodor, J.A. and Pylyshyn, Z.W. 1988. Connectionism and cognitive architecture: A critical analysis. *Cognition* 28, 1–71.

Hameroff, S.R. 1987. *Ultimate Computing: Biomolecular Consciousness and Nanotechnology*. North Holland.

Hofstadter, D.R. 1979. *Gödel, Escher, Bach: An Eternal Golden Braid*. Harvester Press.

Lucas, J.R. 1961. Minds, machines and Gödel. *Philosophy* 36, 120–124. Reprinted in *Minds and Machines,* A.R. Anderson (ed), 1964. Prentice Hall.

Menneer, T. and Narayanan, A. 1995. Quantum-inspired neural networks. Research Report 329, Department of Compute Science, University of Exeter.

Newell, A. and Simon, S.A. 1976. Computer science as empirical enquiry. *Communications of the ACM* 19, 113–126. Also in *Mind Design,* J. Haugeland (ed), 1981, 35–66. Cambridge, MA: MIT Press.

Niklasson, L.F. and Sharkey, N. 1994. Connectionism — the miracle mind model. In *Connectionism in a Broad Perspective*, L.F. Niklasson and M.B. Bodén (eds), 13–25. Ellis Horwood.

Niklasson, L.F. and van Gelder, T. 1994. On being systematically connectionist. *Mind and Language* 9(3), 288–302.

Penrose, R. 1989. *The Emperor's New Mind.* Oxford University Press.

Penrose, R. 1994. *Shadows of the Mind.* Oxford University Press.

Pribram, K.H. 1986. The cognitive revolution and mind/brain issues. *American Psychologist* May, 507–520.

Putnam, H. 1967. The nature of mental states. In *Mind and Cognition*, W.G. Lycan (ed), 1990, 47–56. Blackwell.

Rapaport, W.J. 1988. Syntactic semantics: Foundations of computational natural language understanding. In *Aspects of Artificial Intelligence*, J.F. Fetzer (ed). Kluwer Academic Press.

Rumelhart, D.E. and McClelland, J.L. 1985. Levels indeed! A response to Broadbent. *Journal of Experimental Psychology: General* 114(2), 193–197.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition — Volume 1: Foundations*. Cambridge, MA: MIT Press.

Searle, J.R. 1980. Minds, brains and programs. *The Behavioral and Brain Science* 3, 63–73.

Searle, J.R. 1987. Minds and brains without programs. In *Mindwaves*, C. Blakemore and S. Greenfield (eds). Blackwell.

Turing, A.M. 1950. Computing machinery and intelligence. *Mind* LIX 433–460.

van Gelder, T. 1990. Compositionality: A connectionist variation on a classical theme. *Cognitive Science* 14, 355–364.

von Neumann, J. 1966. *Theory of self-reproducing automata*. Edited and completed by A.W. Banks. University of Illinois Press.

Winfree, A.T. and Strogatz, S.H. 1984. Organizing centres for three dimensional chemical waves. *Nature* 311, 611–615.

The Search for Mind

A New Foundation for Cognitive Science

Seán Ó Nualláin
Dublin City University and NRC, Canada

1. Cognitive Science in Crisis

Paradoxically, it is a compliment to a young science such as Cognitive Science to state that it is in crisis. Kuhn's (1962) justly celebrated account of scientific discovery, using Heidegger's earlier terminology, distinguishes between periods of "normal" science and periods of crisis when fault-lines in the foundations of the discipline offer an opportunity for a paradigm-change, i.e., a change in its basic concepts which is irreversible. A great deal of progress in the discipline is necessary before its fundamental tenets become clear enough to be seen as being out of kilter with the field it claims to study. The concern of this section is to establish that, at least on an "as if" basis (Kuhn, we shall see, has his critics), such is the case for Cognitive Science, considered as the Science of Mind, at this moment.

What seemed at one point a naturally-occurring (super-) discipline with its subject a natural domain (i.e. "informavores"; see Pylyshyn 1984: xi) is now under attack on both theoretical and empirical grounds, with objections both to its methodological and substantive aspects (it would perhaps be relevant to point also to flaws in its pretheoretic specification of the domain, which we discuss below; however, that is more a philosophy of science issue). Let us very briefly examine the picture of Cognitive Science which is under attack, lest the assault be an attention-seeking fracas with a straw man. The domain of Cognitive Science is agreed as perception and knowledge, broadly defined (*ibid.* and Gardner 1985: 6); its tenets are the acceptance of a level of representation, the use of computers, de-emphasis (now nuanced) on social factors and affect, and an interdisciplinary ethos which is the only currently uncontroversial point. Usually, the only reason that materialism is not mentioned as a tenet is that it is regarded as *a priori* valid. Of these tenets, only the interdisciplinary ethos is likely to survive the decade.

Pylyshyn (*op. cit.*), with considerable ingenuity, attempted truly to found a discipline on these tenets. It might ignore consciousness and affect; it might take as its paradigmatic example of mind-world interaction the rarefied universe suggested by the interpretation of states in registers as binary numbers; however, apart from Von Eckardt (1993) which we discuss below, it remains the best worked-out foundation for computationalist Cognitive Science. It is in the epistemological assumptions inherent in Pylyshyn's model that we come across the most damaging attack on it. Edelman (1992: 230–234) argues persuasively that it contains many unproven assumptions about the structure of the world *and* the way we categorize it. Pylyshyn's Cognitive Science requires that a domain can be represented in a fashion which yields itself to syntactic analysis followed by semantic interpretation like the example involving states in registers he chooses. Edelman argues, correctly in my view, that this requires that the world be structured in classical categories, and that this structure informs our perception of it. (Searle 1992, echoing Kripke on Wittgenstein, is unprepared to admit that syntax is intrinsic to the physics; a human agent is necessary for the syntactic analysis as for the semantic interpretation). Rosch (1973) indicates that the notion that our perception of the world is "classical" is not true; the other point is perhaps better stated as "there exists a neat semantic description s for each cognitive domain c" which is also untrue. Wittgenstein's (1922, 1967) path from the neat to the scruffy camp is perhaps the most instructive demonstration of its untruth.

Edelman (*ibid.*) continues his epistemological attack on computationalist Cognitive Science with the argument that it requires that the world be in some non-trivial sense an infinite Turing tape, which of course it isn't. Moreover, Edelman argues, only with continual reference to evolutionary history and consequent neurological structure will any cognitive process be elucidated. However, his knowledge of many of the other constituent disciplines of Cognitive Science is very incomplete, showing the need for the administrative structures of the discipline whose conceptual foundations he has been so keen to criticize; Edelman would greatly benefit from an unbiased reading of Gardner.

Edelman's position on meaning, i.e. the insistence that it requires consciousness and embodiment, is a distant echo of the work of Michael Polanyi (1958) and all the more secure for this resemblance. The accent on consciousness is expanded on by Searle (1992), who finds it scandalous that a science allegedly concerned with Mind should ignore its conscious aspect. Moreover, Searle insisted that neither materialism or dualism is a tenable, or indeed a coherent position. However, Searle (1992: 228) is after bigger game; the limitation of inquiry in the sciences of mind to two levels i.e. the neurophysiological and the phenomenological. This is in explicit contrast to the central tenet of Cognitive Science, as proposed by Dennett (1993: 195):

> There is a level of analysis, the information-processing level, intermediate between the phenomenological level and the neurophysiological level.

Several good reasons exist for preferring Dennett's account of this particular event to Searle's. The first is that, having abandoned one level of analysis, there does not *a priori* seem to be any good principled reason for not abandoning others. For example, the neurophysiological level can successively be reduced to levels in which the explanatory frameworks of chemistry and sub-atomic physics are the most relevant. Secondly, Dennett's central tenet does not strongly constrain Cognitive Science; it can be interpreted as signalling that such a level is interesting and important without requiring one to buy into the notion that it is intentional and cashing out consciousness in this way. Thus interpreted, this tenet sits easily with the viewpoint on the study of mind outlined below.

The cognitive role of emotion is treated by de Sousa (1987), who insists that the fundamental role of emotion is to compensate for the insufficiency of reasoning, manifest *in excelsis* in the Frame problem (*op. cit.*, 195). Of the original pillars of AI, that leaves only representation standing. The work of Gibson (1979) seemed to have safely been dismissed (e.g. Pylyshyn 1984) until the successful use of its findings in computer vision, *inter alia*; Brook's (1991) provocative attacks on centralized symbolic representation forced the most diehard "Establishment" members at least to re-examine certain basic concepts.

The most recent thorough attempt to found Cognitive Science on classical lines is due to Von Eckardt (1993). Her argument on Cognitive Science is in itself powerful enough to merit attention; however, she explicitly denies that Cognitive Science is in Kuhnian crisis, but rather is an immature Science with an implicit set of commitments (30–31) which her book succeeds in making explicit (13).

Immature science is exemplified for Von Eckardt (353) by notions like the fluid theory of electricity. A paradigm change, on the other hand, could be provoked by a phenomenon like black-body radiation. Let's imagine that Physics had continued to ignore or explain away black-body radiation, and had continued in its long-accustomed path; such was the case for the Ptolemaic Universe (see Aizawa's paper in the next section). Is this self-blinkering not precisely analogous to the current attempts of Cognitive Science to establish a science of Mind without consciousness, emotion and social factors? What we have is an almost