ARTIFICIAL INTELLIGENCE

Its Philosophy and Neural Context

F.H. George

ROUTLEDGE LIBRARY EDITIONS:
ARTIFICIAL INTELLIGENCE

ROUTLEDGE

ROUTLEDGE LIBRARY EDITIONS:
ARTIFICIAL INTELLIGENCE

Volume 5

ARTIFICIAL INTELLIGENCE

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

ARTIFICIAL INTELLIGENCE
Its Philosophy and Neural Context

F.H. GEORGE

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published in 1986 by Gordon and Breach Science Publishers

This edition first published in 2018
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-0-8153-8566-0 (Set)
ISBN: 978-0-429-49236-5 (Set) (ebk)
ISBN: 978-0-8153-6336-1 (Volume 5) (hbk)
ISBN: 978-1-351-10959-8 (Volume 5) (ebk)

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original copies may be apparent.

Disclaimer
The publisher has made every effort to trace copyright holders and would welcome correspondence from those they have been unable to trace.

Artificial Intelligence
Its Philosophy and Neural Context

F.H. GEORGE
Brunel University, U.K.

GORDON AND BREACH SCIENCE PUBLISHERS
New York London Paris Montreux Tokyo

Gordon and Breach Science Publishers

P.O. Box 786
Cooper Station
New York, NY 10276
United States of America

P.O. Box 197
London WC2E 9PX
England

58, rue Lhomond
75005 Paris
France

14-9 Okubo 3-chome
Shinjuku-ku
Tokyo 160
Japan

Library of Congress Cataloging in Publication Data

George, F.H. (Frank Honywill), 1921–
Artificial intelligence: its philosophy and neural context.

(Studies in cybernetics, ISSN 0275–5807; v. 9)
Bibliography: p.
Includes index.
1. Artificial intelligence. 2. Information science.
I. Title. II. Series.
Q335.G385 1984 001.53′5 84–10253
ISBN 2–88124–009–7 (Switzerland)

ISBN 2–88124–009–7. ISSN 0275 5807. Printed in the United Kingdom by Bell and Bain Ltd, Glasgow.

Introduction to the Series

The subject of cybernetics is quickly growing and there now exists a vast amount of information on all aspects of this broad-based set of disciplines. The phrase "set of disciplines" is intended to imply that cybernetics and all the approaches to artificial (or machine) intelligence have a near identical viewpoint. Furthermore, systems analysis, systems theory and operational research often have a great deal in common with (and are in fact not always discernibly different from) what is meant by cybernetics, as far as this series is concerned: inevitably, computer science is bound to be involved also.

The fields of application are virtually unlimited and applications are discovered in the investigation or modelling of any complex system. The most obvious applications have been in the construction of artificially intelligent systems, the brain and nervous system, and socio-economic systems. This can be achieved through either simulation (copying as exactly as possible) or synthesis (achieving the same or better end result by any means whatsoever).

The range of applications today has become so broad it now includes such subjects as aesthetics, history and architecture. Modelling can be carried out by computer programs, special purpose models (analog, mathematical, statistical, etc.), and automata of various kinds, including neural nets and TOTES. All that is required of the system to be studied is that it be complex, dynamic, and capable of 'learning' and also have feedback or feedforward or both.

This is an international series.

FRANK GEORGE

Contents

Preface

In this book I have attempted to place Artificial Intelligence (AI) in the context of cybernetic thinking and information science. This involves giving some consideration to philosophical, behavioural and neurophysiological matters. It is this contextual issue that is my main interest although at the same time I hope to consider the AI approach to cognition in a fair amount of detail.

I have had the benefit of reading the following important works: Patrick Winston and Richard Brown's *Artificial Intelligence: An MIT Perspective*, Margaret Boden's *Artificial Intelligence and Rational Man*, R.W. Sampson's *Adaptive Information Processing*, the various and considerable writing of Donald Michie, L. Johnston's and E.T. Kervanou's *Expert Systems Technology* and Alex Andrew's *Artificial Intelligence* — as well, of course, as all the books and papers listed in the references. This book is in no way intended to compete with all these excellent books — without which I would know much less about AI; it is though intended to provide a link between AI and the worlds of philosophy, logic, neurophysiology and the like: to provide a sort of context for AI.

Inevitably when writing such a book, especially in a field such as AI which is developing very quickly, there are many different ways one could place the emphasis. I considered saying a great deal more about the tools for building expert systems (expert system builders such as AGE and EMYCIN) and more on Knowledge Elicitation, Knowledge Representation and expert systems themselves. Instead I chose to concentrate on expert systems and in particular on two consultative expert systems — PROSPECTOR and INTERNIST, thanks especially to the careful and painstaking guidance of Drs. Johnston and Kervanou. I am planning in the future to provide a more detailed analysis of expert systems.

The term "expert system" is somewhat ambiguous since the work on GPS by Newell, Shaw and Sinom, Checkers by Samuel and GT by Donald Michie are all expert systems as I see it, but they are not consultative in that they do not interact with the user or explain (justify) the advice given — the title "consultative expert system" is often though identified with "expert system".

I profoundly believe that AI, and rather especially expert systems, are going to play a major part in the future of both science and business and in a way this is something of a triumphant justification for the originators of such thinking — particularly Norbert Wiener and Warren McCulloch.

FRANK GEORGE

BOOK 1

The Basic Issues of Information Technology

Preface to book 1

In Book 1, I shall try to set down what I conceive to be the most central and most basic issues of information technology, which I shall take to include cybernetics and artificial intelligence.

These basic issues will include some discussion of the distinction which is increasingly made by some people between cybernetics and artificial intelligence (AI), although I personally still think of cybernetics as including AI. It is perhaps though a matter of no great importance how we make such classifications.

The other main considerations include some statement, in the form of guidelines, if not precise definitions, as to what is to be regarded as a cybernetic problem; I shall call it this if only for brevity's sake. This leads to a consideration as to which systems can, or should, be studied cybernetically, and also to the realisation that cybernetics is, in some measure, *a point of view*. It is a methodology on its own, has a philosophy of its own, but its systems and their modelling, it takes from wherever it finds them. This entails a considerable overlap with other disciplines both as a source of systems to analyse and models to construct as well as in terms of methodology.

There are many special questions of a philosophical kind that arise and these, in one way or another, surround the mind–body problem problem. We shall examine all these matters in Book 1, and they lead on to a more detailed analysis in Book 2.

1

The Nature of the Problem

In this book I hope to achieve a certain number of reasonably clear-cut goals. The first is to outline what I believe to be the principal problems of cybernetics, especially with respect to artificial intelligence (AI). Herein lies the first important issue.

When Wiener and his associates (1948) first introduced the notions of cybernetics, and Shannon (Shannon and Weaver, 1949) first drew attention to the importance of information theory, and the group of scientists and philosophers in North America in the 1940s first held discussions on philosophy of science and cybernetics, there was an upsurge of interest in "man as a machine". McCulloch and Pitts (1943) in America, and Turing in Britain (1951), among many others from various disciplines, further stimulated that interest. Out of their work came renewed efforts in the modelling of nervous systems and the development of particular modelling methods such as neural nets, leading both to and from automata theory in general and such methods as TOTES in particular.

At the same time as this was happening, there arose a new interest in purposive behaviour and teleological explanation. Furthermore, thanks to the interest of Gregory Bateson, Margaret Mead and others, the ideas of cybernetics were seen to be capable of being used to model social, anthropological, economic, indeed almost any sort of system in which scientists were interested.

There were underlying themes (which we shall discuss later in more detail) regarding feedback and feedforward, adaptation, and so on. All this lay within the general context of the belief that, despite the semantic difficulties involved, there was a sense in which "machines could be made to think".

Since these early days, a great deal has happened. The hugely over-optimistic forecasts that progress would, in the fifties and sixties, be electrifying, did not materialise. This is fairly typical historically, since a "breakthrough" often stirs up enthusiasm and over-statement, however well intentioned. However, my theme here is more concerned with one particular aspect of the subsequent history of cybernetics. What is now regarded as

AI, and workers within their field think of it as separated from (broken away from) cybernetics, did indeed emerge in a separate form. Rather as a liberal party, having done its job of liberalisation, may cease to exist, so cybernetics having persuaded other disciplines to look at themselves "cybernetically" was in danger of disappearing. This disappearance being encouraged by the unwanted presence of a lunatic fringe (much the same thing has happened to psychology).

So in writing now about cybernetics *and* AI, I am conceding that the two have in some measure parted company, even though I shall think of the latter as being largely included in the former. (*pace* AI—folk). In his recent book, Alex Andrew (1983) has quoted me specifically as "identifying cybernetics and AI", and reasonably so.

Historically, the basic idea that "machines" *could* be made to "think" required interpretation and evidence by implementation. This meant following up modelling methods such as neural nets, but it was soon realised that in practice the only universal model capable of realistic use was the suitably programmed computer. Noughts-and-crosses (tic-tac-toe), NIM and other games were quickly tackled and soon came the cascade of ideas and programming methods (LISP, heuristics, etc.) which led to the division of labour between those who programmed computers and those who did other things. "What are those other things?" is the first question. But before we attempt to answer it let it be said that there is an important distinction (of degree no doubt) between the *simulation* of human behaviour (of special interest to psychologists and physiologists) and the *synthesis* of cognitive activities which has great value in the field of control engineering and many other fields besides.

The distinction between simulation and synthesis cuts across the current cybernetic-AI distinction, since many cyberneticans are interested in computer modelling and many AI people are concerned with how their work helps in understanding human behaviour, both "normal" and "abnormal".

If however we admit, however temporarily, the cybernetic-AI distinction, we might ask what is left when the AI work has been subtracted. The answer is (or should be) of special interest, I think, to both parties. The remainder includes all other model making methods such as automata, information theory, TOTES and the application of formalisation. The fields of activity to which these other methods can now be addressed is everything and anything at all (e.g. history, economics, the arts) in just the same way as AI methods could apply to anything at all. Because the "other" methods are less easily demarcated, their use can often be buried in the field to which they are applied. It does appear though that this is now happening to AI as well.

The actual situation that then pertains is that all too often AI people will regard cybernetics as an empty subject. This picture, which is characteristic of the sort of absurd extremes that arise in science and philosophy, is wholly ridiculous. It is reminiscent of Russell's outright condemnation of Wittgenstein's second phase philosophy and Wittgensten's subsequent loathing of *Principia Mathematica* and the formalism it stood for. Such formalism was baptised the Augustinian viewpoint (Baker and Hacker, 1980) and presented as the rigid and the backward looking. Such feuds abound and are doubtless the result of fervant enthusiasm and dedication, but none the less fatuous for all that.

For my own part I think of AI as part of the spearhead of cybernetics and a field of immense potential. Indeed in this book I would lay no claim to describing AI in and of itself, since this has already been excellently done by many writers including Margaret Bowden, E. Charniack and Y. Wilks, R. Sampson, Winston and Brown, Nilsson, Slagle and many others. These are works mainly by people currently active in the field of AI and I would not for a moment suggest this book is an attempt to compete with their work. Similarly it is no function of any book, least of all this one, to attempt to compete with those two excellent sources of papers on AI, the series of books on *Machine Intelligence*, from Edinburgh and edited by Donald Michie and his associates, and the equally excellent journal *Artificial Intelligence*, edited by Daniel Bobrow; there are also other journals that are springing up like mushrooms. I have drawn extensively on all these sources to try to give a picture of AI at work, in the general context of cybernetic thinking.

If cybernetics is concerned with precision and effectiveness in modelling complex, interactive adaptive feedback and feedforward systems, this in practice will entail supplying a formalised account of fields of activity as varied as history (Chandler 1984) and art (Rosenberg 1983) or business science (Strank 1983). I believe this sort of approach entails looking carefully at the philosophical and semantic aspects of the development; the behavioural fields could also entail looking in turn at the neurophysiological evidence. I hope to do just that, but only with a view to providing guidelines to AI work in their simulatory role.

Marr's work (1974, 1975, 1982) in characteristic of those AI researchers whose primary interest in simulation of the nervous system. He saw visual sensation and visual perception as starting with the formation of a "primal sketch" which is rather like the primitive and important features of a scene. These, for Marr, include shading-edge, extended edge, line, contrast, lighters, size, etc. All this is in the periphery of the visual system and therefore primarily retinal (i.e. mechanisms of detection).

This is followed by the central processes (mechanism of construction) which group and organise (e.g. prints) figures-ground data and generally provide an interpretation of the primal sketch. This is similar to (but not quite the same as) Price's (1953) distinction between primary and secondary recognition. In fact, it is rather like the distinction sometimes made between *sensing* (visually) and *perception.* But in any case there may be many stages which are more or less separable in the total visual activity.

The great virtue, or so it seems to me, is that AI provides a powerful tool which goes beyond traditional experimental psychology and neuropsychological modelling and theorising (e.g. Hebb, 1949; Palm, 1982) and this should be recognised by all those parties interested in the function of the senses. Dr Marr's (1975) own words are worth quoting here:

> The situation in modern neurophysiology is that people are trying to understand how a particular mechanism performs a computation that they cannot even formulate, let alone provide a crisp summary of ways of doing so. To rectify the situation, we need to invest considerable effort in studying the computational backgrounds to questions that can be approached in neurophysiological experiments.

He goes on to emphasise, quite rightly, that this work is an indirect approach to neurophysiology, but a direct approach to vision. This is the synthesis coming first, but with the simulation in mind and with a thorough awareness of the great complexity involved. He makes the point that until you have tried to suitably program a computer to perform even relatively simple looking tasks you cannot have any feel for the complexity of the problem. His words again:

> The power of this approach is that the knowledge one obtains concerns facts that are inherent in the task, not in the structural details of the mechanism performing it.

With all of this, I would readily and unconditionally agree, but would like to add a caveat. It is all too characteristic of philosophers that they will support a viewpoint: Wittgenstein's constructivism, Peirce's pragmatism, Russell's atomism or Ryle's natural language approach, and see them less as alternatives and acceptable (even if not equally acceptable) approaches to the philosophy of language, say, and more as the one single correct approach (theirs) and the others as incorrect approaches (all those that are not theirs). This same attitude has often pervaded science and one remembers experimental psychologists of the 1930s, 1940s and 1950s, with their love of apparatus and statistical analysis, and their loathing of theory. The danger is that the enthusiasm of AI people could similarly carry them so far that they could see absolutely no virtue at all in any other approach. For example, to take a philosophical approach and try to carefully analyse one's language, in order among other things, to try to decide whether a problem is a "real" one or not. So the nativist-empiricist controversy in perception was wasteful and absurd since it was a totally semantic issue and could be resolved by almost immediately a little linguistic — philosophical consideration. It was time wasted when one might be writing and testing programs.

It is not being suggested that AI people see it in this extreme light any more than cyberneticians outside AI see AI people as wasting their time. There are doubtless minority groups in both camps though who do think this, but it would be sad to see this parochiality go too far. Another example which is closely parallel is the feelings that often exist between logicians — mathematical on one hand and non-mathematical on the other. That there are various possible approaches to logic is obvious but it is not easy for people wholly involved in the more extreme removes of axiomatics and recursive functions to see much use in the "limpid world" (Quine's words) of those like Strawson or Austin.

It has also been said against the AI group that they claim for their work a case for simulation that could never in fact go beyond synthesis. I know of no case personally where excessive claims have been made but I would guess that the vast majority of people in the field must know that as long as we are in doubt about the way human mechanisms function, we cannot be sure that we possess a precise simulation model. It seems virtually certain in cases such as visual perception that template models, PERCEPTRON-like models and the various neural models cannot all be simulations. Similarly in syntactical analysis augmented-transition nets, Marcus type syntax and forms of dependency grammar cannot all be simulations of the way human beings deal with language. Perhaps none of them are simulations or perhaps some two or three are part of a more general simulation; all of this must, at least for the time being, remain uncertain.

I have though, returning to the suggested AI-cibernetics distinction, seen it said (a private communication from a distinguished AI worker) that cybernetics, while worthy enough when originally founded by Wiener and Shannon, "had died but simply refused to lie down". The answer here is that it has changed and evolved and much has rubbed off on other disciplines and been incorporated by those other disciplines, and if you think of AI, not as the central thrust of cybernetics, as I do, but as something now working on its own, there is still a great deal remaining in cybernetics, including the various methods and the application of those methods to any field of activity, but perhaps especially to human behaviour and the human nervous system. From this, I believe, a very valuable association with philosophy is growing up, where the ambitions of modern pragmaticists (e.g. Montague, Stalnaker and Martin), people in cognitive linguistics (e.g. Osgood and Rozeboom) and meaning theorists, (e.g. Quine, Bennett and Grice) can find a valuable trade off. The cybernetician should be greatly involved in such linguistic, logical and behavioural undertakings and add some emphasis to them. At the same time his own models should be built on them to put-it-another-way, the field of overlap here, perhaps the field of (descriptive) pragmatics is part of cybernetics and its models — gradually appearing in automata theoretic or program form — is a bridge to AI.

For my own part I am personally concerned (apart from my work in AI) in research is what is perhaps best called behavioural semiotic. I think of this as a valid field for a cybernetician to work in, but would add that whether or not you call this (or *could* call this) cybernetics is of absolutely no consequence. This raises that other chestnut about the (historically) useful but often misleading distinctions that occur (or can be thought to occur) between disciplines so that one asks "What is a philosophical problem?" "Is that a part of physics or chemstry?" and so on. One reason for the founding of cybernetics, rather like operational research, was to cut across the more obviously arbitrary boundaries of existing success and "slice the cake" differently. It is partly for this reason that we now have so many labels and to tilt at the labels (and their roughly delineated adherents) without too much consideration of the detailed work of some of the people, regardless of the labels, seems foolish.

Trying to classify what AI people do, what cyberneticians do, and their relation to each other is one clearcut aim, another is to try to outline the methods of modelling which are available to scientists — primarily information scientists, as well as outline and evaluate progress both in the wider (cybernetic) and also the narrower (strictly AI) approaches.

I would like also to say something about the present controversy (Searle, 1982) over "strong" and "weak" viewpoints in AI. Searle seems to argue that a simulation (or model) of a human being (or anything at all) is different from the "thing" (human being) itself. A model that is a "complete" duplicate is hardly appropriately called a model. The trouble with these sorts of discussions is that one is often dealing with a "straw man": no one seems to hold the view they are supposed to. But we can at least point to a subsidiary question as to whether the fabric of a system directly affects, or not, the behaviour of a system. This is a familiar argument for which there are no definite answers available. But it comes to this: can an artifically constructed system have exactly the same experiences as a human being?

The traditional philosophical problem of "other minds" rears its head here because there is no way we can be certain that *other people* have the same feelings as we have. It is constantly the case, as far as even cognition is concerned, let alone conation, that we cannot seriously encompass our answers in the Turing Interrogation Game, since the constraints imposed by it are no more than suggestive of the criteria we need. In fact, we can never *know* (if that implies indubitibility) whether "machines" (as defined as "artificially constructed systems") will have the property of consciousness, but we might guess that if they get to the point of holding humanlike conversation with self-reference, etc., then consciousness may prove to be inevitable. It seems likely that for our model to be humanlike in all respects may require that it be made of colloidal protoplasm and be exposed to a similar range of experiences. But this is too fanciful for the moment. What one might contend is that there is no obvious reason why such a state of affairs might not be achieved. However, on the short term, all people interested in AI will be satisfied if we continue to accumulate more and more information about humanlike systems and their behaviour, and continue to compare and contrast the consequent models with the human to see when they are apparently getting nearer to an understanding of the principles upon which the human being functions.

There are a number of people — some even within cybernetics and AI — who are sceptical or appalled, or both, by what is sometimes claimed, or what is anyway implicit in the whole project. Hubert Dreyfus (1965, 1972) stands out as one of these who believes that far too much is claimed for the computer. It does not seem worthwhile to enter in any detail into the argument at this point, since Margaret Boden has adequately dealt with many of the misunderstandings embodied in Dreyfus' other objections such as those which stem from Gödel's theorems. In general, the limitations that apply to the abilities of computers are limitations that apply in much the same way to human beings, and although one should be careful about claims made on behalf of work in cybernetics and AI, most of the objections put forward, in such a form as Dreyfus' arguments, are usually simply incorrect.

The scene is now set in part I of this book to look at cybernetics and modelling methods with some reference to philosophy, neuro-physiology and psychology and then to explicitly discuss AI in part II in terms of the same sort of behavioural context.