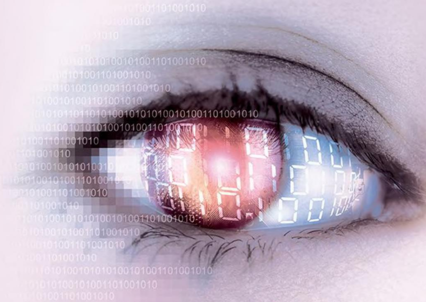
MARGARET A. BODEN



Its nature and future

AI

AI

Its nature and future

MARGARET A. BODEN

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS
Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

First Edition published in 2016

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number
Data available

ISBN 978–0–19–877798–4

Printed in Great Britain by
Clays Ltd, St Ives plc

CONTENTS

LIST OF FIGURES

1

What is Artificial Intelligence?

Artificial intelligence (AI) seeks to make computers do the sorts of things that minds can do.

Some of these (e.g. reasoning) are normally described as "intelligent." Others (e.g. vision) aren't. But all involve psychological skills—such as perception, association, prediction, planning, motor control—that enable humans and animals to attain their goals.

Intelligence isn't a single dimension, but a richly structured space of diverse information-processing capacities. Accordingly, AI uses many different techniques, addressing many different tasks.

And it's everywhere.

AI's practical applications are found in the home, the car (and the driverless car), the office, the bank, the hospital, the sky… and the Internet, including the Internet of Things (which connects the ever-multiplying physical sensors in our gadgets, clothes, and environments). Some lie outside our planet: robots sent to the Moon and Mars, or satellites orbiting in space. Hollywood animations, video and computer games, sat-nav systems, and Google's search engine are all based on AI techniques. So are the systems used by financiers to predict movements on the stock market, and by national governments to help guide policy decisions in health and transport. So are the apps on mobile phones. Add avatars in virtual reality, and the toe-in-the-water models of emotion developed for "companion" robots. Even art galleries use AI—on their websites, and also in exhibitions of computer art. Less happily, military drones roam today's battlefields—but, thankfully, robot mine-sweepers do so too.

AI has two main aims. One is *technological*: using computers to get useful things done (sometimes by employing methods very *unlike* those used by minds). The other is *scientific*: using AI concepts and models to help answer questions about human beings and other living things. Most AI workers focus on only one of these, but some consider both.

Besides providing countless technological gizmos, AI has deeply influenced the life sciences. A computer model of a scientific theory is a test of its clarity and coherence, and a compelling demonstration of its—often unknown—implications. Whether the theory is *true* is another matter, and depends on evidence drawn from the science concerned. But even discovering that it's false can be illuminating.

In particular, AI has enabled psychologists and neuroscientists to develop powerful theories of the mind-brain. These include models of *how the physical brain works,* and—a different, but equally important, question—*just what it is that the brain is doing*: what computational (psychological) questions it is answering, and what sorts of information processing enable it to do that. Many unanswered questions remain, for AI itself has taught us that our minds are very much richer than psychologists had previously imagined.

Biologists, too, have used AI—in the form of "artificial life" (A-Life), which develops computer models of differing aspects of living organisms. This helps them to explain various types of animal behavior, the development of bodily form, biological evolution, and the nature of life itself.

Besides affecting the life sciences, AI has influenced philosophy. Many philosophers today base their accounts of mind on AI concepts. They use these to address, for instance, the notorious mind–body problem, the conundrum of free will, and the many puzzles regarding consciousness. However, these philosophical ideas are hugely controversial. And there are deep disagreements about whether any AI system could possess *real* intelligence, creativity, or life.

Last, but not least, AI has challenged the ways in which we think about humanity—and its future. Indeed, some people worry about whether we actually have a future, because they foresee AI surpassing human intelligence across the board. Although a few thinkers welcome this prospect, most dread it: what place will remain, they ask, for human dignity and responsibility?

All these issues are explored in the following chapters.

Virtual Machines

"To think about AI," someone might say, "is to think about computers." Well, yes and no. The computers, as such, aren't the point. It's what they *do* that matters. In other words, although AI needs *physical* machines (i.e. computers), it's best thought of as using what computer scientists call *virtual* machines.

A virtual machine isn't a machine depicted in virtual reality, nor something like a simulated car engine used to train mechanics. Rather, it's the *information-processing system* that the programmer has in mind when writing a program, and that people have in mind when using it.

As an analogy, think of an orchestra. The instruments have to work. Wood, metal, leather, and cat-gut all have to follow the laws of physics if the music is to happen as it should. But the concert-goers aren't focused on that. Rather, they're interested in the music. Nor are they concerned with individual notes—still less, with the vibrations in the air that are causing the sound. They're listening to the musical "shapes" made up by the notes: the melodies and harmonies, themes and variations, slurs and syncopation.

Where AI is concerned, the situation is similar. A word processor, for example, is thought of by its designer, and experienced by its users, as dealing directly with words and paragraphs. But the program itself usually contains neither. (Some do, e.g. copyright notices, which can be easily inserted by the user.) And a neural network (see Chapter 4) is thought of as doing information processing *in parallel*, even though it's usually implemented in a (sequential) von Neumann computer.

That's not to say that a virtual machine is just a convenient fiction, a thing merely of our imagination. Virtual machines are actual realities. They can make things happen, both inside the system and (if linked to physical devices such as cameras or robot hands) in the outside world. AI workers trying to discover what's going wrong when a program does something unexpected only rarely consider hardware faults. Usually, they're interested in the events and causal interactions in the *virtual* machinery, or software.

Programming languages, too, are virtual machines (whose instructions have to be translated into machine code before they can be run). Some are defined in terms of lower-level programming languages, so translation is required at several levels. They're needed because most people can't think about information processing in the bit patterns used for machine code, and no one can think about complex processes at that hugely detailed level.

That's not true only of programming languages. Virtual machines in general are comprised of patterns of activity (information processing) that exist at various levels. Moreover, it's not true only of virtual machines running on computers. We'll see in Chapter 6 that *the human mind* can be understood as the virtual machine—or rather, the set of mutually interacting virtual machines, running in parallel (and developed or learned at different times)—that is implemented in the brain.

Progress in AI requires progress in defining interesting/useful virtual machines. More *physically* powerful computers (larger, faster) are all very well. They may even be necessary for certain kinds of virtual machines to be implemented. But they can't be exploited unless *informationally* powerful virtual machines can be run on them. (Similarly, progress in neuroscience requires better understanding of what *psychological* virtual machines are being implemented by the physical neurons: see Chapter 7.)

Different sorts of external-world information are used. Every AI system needs input and output devices, if only a keyboard and a screen. Often, there are also special-purpose sensors (perhaps cameras, or pressure-sensitive whiskers) and/or effectors (perhaps sound synthesizers for music or speech, or robot hands). The AI program connects with—causes changes in—these computer-world interfaces as well as processing information internally.

AI processing usually also involves *internal* input and output devices, enabling the various virtual machines within the whole system to interact with each other. For example, one part of a chess program may detect a possible threat by noticing something happening in another, and may then interface with yet another in searching for a blocking move.

The Major Types of AI

How the information is processed depends on the virtual machine involved. As we'll see in later chapters, there are five major types, each including many variations. One is classical, or symbolic, AI—sometimes called GOFAI (Good Old-Fashioned AI). Another is artificial neural networks, or connectionism. In addition, there are evolutionary programming; cellular automata; and dynamical systems.

Individual researchers often use only one method, but *hybrid* virtual machines also occur. For instance, a theory of human action that switches continually between symbolic and connectionist processing is mentioned in Chapter 4. (This explains why, and how, it is that someone may be distracted from following through on a planned task by noticing something unrelated to it in the environment.) And a sensorimotor device that combines "situated" robotics, neural networks, and evolutionary programming is described in Chapter 5. (This device helps a robot to find its way "home" by using a cardboard triangle as a landmark.)

Besides their practical applications, these approaches can illuminate mind, behavior, and life. Neural networks are helpful for modeling aspects of the brain, and for doing pattern-recognition and learning. Classical AI (especially when combined with statistics) can model learning too, and also planning and reasoning. Evolutionary programming throws light on biological evolution and brain development. Cellular automata and dynamical systems can be used to model development in living organisms. Some methodologies are closer to biology than to psychology, and some are closer to non-reflective behavior than to deliberative thought. To understand the full range of mentality, all of them will be needed—and probably more.

Many AI researchers don't care about how minds work: they seek technological efficiency, not scientific understanding. Even if their techniques originated in psychology, they now bear scant relation to it. We'll see, however, that progress in general-purpose AI (artificial general intelligence, or AGI) will require deep understanding of the computational architecture of minds.

AI Foreseen

AI was foreseen in the 1840s by Lady Ada Lovelace.[1] More accurately, she foresaw *part* of it. She focused on symbols and logic, having no glimmering of neural networks, nor of evolutionary and dynamical AI. Nor did she have any leanings towards AI's psychological aim, her interest being purely technological.

She said, for instance, that a machine "might compose elaborate and scientific pieces of music of any degree of complexity or extent," and might also express "the great facts of the natural world" in enabling "a glorious epoch in the history of the sciences." (So she wouldn't have been surprised to see that, two centuries later, scientists are using "Big Data" and specially crafted programming tricks to advance knowledge in genetics, pharmacology, epidemiology… the list is endless.)

The machine she had in mind was the Analytical Engine. This gears-and-cogwheels device (never fully built) had been designed by her close friend Charles Babbage in 1834. Despite being dedicated to algebra and numbers, it was essentially equivalent to a general-purpose digital computer.

Ada Lovelace recognized the potential generality of the Engine, its ability to process symbols representing "all subjects in the universe." She also described various basics of modern programming: stored programs, hierarchically nested subroutines, addressing, microprogramming, looping, conditionals, comments, and even bugs. But she said nothing about *just how* musical composition, or scientific reasoning, could be implemented on Babbage's machine. AI was possible, yes—but how to achieve it was still a mystery.

How AI Began

That mystery was clarified a century later by Alan Turing. In 1936, Turing showed that every possible computation can in principle be performed by a mathematical system now called a universal Turing machine.[2] This imaginary system builds, and modifies, combinations of binary symbols--represented as "0" and "1." After codebreaking at Bletchley Park during World War II, he spent the rest of the 1940s thinking about how the abstractly defined Turing machine could be approximated by a physical machine (he helped design the first modern computer, completed in Manchester in 1948), and how such a contraption could be induced to perform intelligently.

Unlike Ada Lovelace, Turing accepted both goals of AI. He wanted the new machines to do useful things normally said to require intelligence (perhaps by using highly unnatural techniques), and also to model the processes occurring in biologically based minds.

The 1950 paper in which he jokily proposed the Turing Test (see Chapter 6) was primarily intended as a manifesto for AI.[3] (A fuller version had been written soon after the war, but the Official Secrets Act prevented publication.) It identified key questions about the information processing involved in intelligence (game playing, perception, language, and learning), giving tantalizing hints about what had already been achieved. (Only "hints", because the work at Bletchley Park was still top-secret.) It even suggested computational approaches—such as neural networks and evolutionary computing—that became prominent only much later. But the mystery was still far from dispelled. These were highly general remarks: programmatic, not programs.

Turing's conviction that AI must be somehow possible was bolstered in the early 1940s by the neurologist/psychiatrist Warren McCulloch and the mathematician Walter Pitts. In their paper "A Logical Calculus of the Ideas Immanent in Nervous Activity,"[4] they united Turing's work with two other exciting items (both dating from the early twentieth century): Bertrand Russell's propositional logic and Charles Sherrington's theory of neural synapses.

The key point about propositional logic is that it's binary. Every sentence (also called a *proposition*) is assumed to be either *true* or *false*. There's no middle way, no recognition of uncertainty or probability. Only two "truth-values" are allowed, namely *true* and *false*.

Moreover, complex propositions are built, and deductive arguments are carried out, by using logical operators (such as *and*, *or*, and *if–then*) whose meanings are defined in terms of the truth/falsity of the component propositions. For instance, if two (or more) propositions are linked by *and*, it's assumed that both/all of them are true. So "Mary married Tom and Flossie married Peter" is true if, and only if, *both* "Mary married Tom" and "Flossie married Peter" are true. If, in fact, Flossie did not marry Peter, then the complex proposition containing "and" is itself false.

Russell and Sherrington could be brought together by McCulloch and Pitts because they had both described binary systems. The *true/false* values of logic were mapped onto the *on/off* activity of brain cells and the *0/1* of individual states in Turing machines. Neurons were believed by Sherrington to be not only strictly on/off, but also to have fixed thresholds. So logic gates (computing *and*, *or*, and *not*) were defined as tiny neural nets, which could be interconnected to represent highly complex propositions. Anything that could be stated in propositional logic could be computed by some neural network, and by some Turing machine.

In brief, neurophysiology, logic, and computation were bundled together—and psychology came along too. McCulloch and Pitts believed (as many philosophers then did) that natural language boils down, in essence, to logic. So all reasoning and opinion, from scientific argument to schizophrenic delusions, was grist for their theoretical mill. They foresaw a time when, for the whole of psychology, "specification of the [neural] net would contribute all that could be achieved in that field."

The core implication was clear: *one and the same theoretical approach*—namely, Turing computation—could be applied to human and machine intelligence. (The McCulloch/Pitts paper even influenced computer design. John von Neumann, then intending to use decimal code, was alerted to it and switched to binary instead.)

Turing, of course, agreed. But he couldn't take AI much further: the technology available was too primitive. In the mid-1950s, however, more powerful and/or easily usable machines were developed. "Easily usable," here, doesn't mean that it was easier to push the computer's buttons, or to wheel it across the room. Rather, it means that it was easier to define new *virtual* machines (e.g. programming languages), which could be more easily used to define higher-level virtual machines (e.g. programs to do mathematics, or planning).

Symbolic AI research, broadly in the spirit of Turing's manifesto, commenced on both sides of the Atlantic. One late-1950s landmark was Arthur Samuel's checkers (draughts) player, which made newspaper headlines because it learned to beat Samuel himself.[5] That was an intimation that computers might one day develop *super*human intelligence, outstripping the capacities of their programmers.

The second such intimation also occurred in the late 1950s, when the Logic Theory Machine not only proved eighteen of Russell's key logical theorems, but found a more elegant proof of one of them.[6] This was truly impressive. Whereas Samuel was only a mediocre checkers player, Russell was a world-leading logician. (Russell himself was delighted by this achievement, but the *Journal of Symbolic Logic* refused to publish a paper with a computer program named as an author, especially as it hadn't proved a *new* theorem.)

The Logic Theory Machine was soon outdone by the General Problem Solver (GPS)[7]—"outdone" not in the sense that GPS could surpass yet more towering geniuses, but in the sense that it wasn't limited to only one field. As the name suggests, GPS could be applied to any problem that could be represented (as explained in Chapter 2) in terms of goals, sub-goals, actions, and operators. It was up to the programmers to identify the goals, actions, and operators relevant for any specific field. But once that had been done, the *reasoning* could be left to the program.

GPS managed to solve the "missionaries-and-cannibals" problem, for example. (*Three missionaries and three cannibals on one side of a river; a boat big enough for two people; how can everyone cross the river, without cannibals ever outnumbering missionaries?*) That's difficult even for humans, because it requires one to go backwards in order to go forwards. (Try it, using pennies!)

The Logic Theory Machine and GPS were early examples of GOFAI. They are now "old-fashioned," to be sure. But they were also "good," for they pioneered the use of *heuristics* and *planning*—both of which are hugely important in AI today (see Chapter 2).

GOFAI wasn't the only type of AI to be inspired by the "Logical Calculus" paper. Connectionism, too, was encouraged by it. In the 1950s, networks of McCulloch-Pitts logical neurons, either purpose-built or simulated on digital computers, were used (by Albert Uttley, for instance[8]) to model associative learning and conditioned reflexes. (Unlike today's neural networks, these did *local*, not *distributed*, processing: see Chapter 4.)

But early network modeling wasn't wholly dominated by neuro-logic. The systems implemented (in analogue computers) by Raymond Beurle in the mid-1950s were very different.[9] Instead of carefully designed networks of logic gates, he started from two-dimensional arrays of randomly connected, and varying-threshold, units. He saw neural self-organization as due to dynamical waves of activation—building, spreading, persisting, dying, and sometimes interacting.

As Beurle realized, to say that psychological processes could be *modeled by* a logic-chopping machine wasn't to say that the brain *actually is* such a machine. McCulloch and Pitts had already pointed this out. Only four years after their first groundbreaking paper, they had published another one arguing that thermodynamics is closer than logic to the functioning of the brain.[10] Logic gave way to statistics, single units to collectivities, and deterministic purity to probabilistic noise.

In other words, they had described what's now called distributed, error-tolerant computing (see Chapter 4). They saw this new approach as an "extension" of their previous one, not a contradiction of it. But it was more biologically realistic.

Cybernetics

McCulloch's influence on early AI went even further than GOFAI and connectionism. His knowledge of neurology as well as logic made him an inspiring leader in the budding cybernetics movement of the 1940s.

The cyberneticians focused on biological self-organization. This covered various kinds of adaptation and metabolism, including autonomous thought and motor behavior as well as (neuro) physiological regulation. Their central concept was "circular causation," or feedback. And a key concern was teleology, or purposiveness. These ideas were closely related, for feedback depended on goal differences: the current distance from the goal was used to guide the next step.

Norbert Wiener (who designed anti-ballistic missiles during the war) named the movement in 1948, defining it as "the study of control and communication in the animal and the machine."[11] Those cyberneticians who did computer modeling often drew inspiration from control engineering and analogue computers rather than logic and digital computing. However, the distinction wasn't clear-cut. For instance, goal differences were used both to control guided missiles and to direct symbolic problem solving. Moreover, Turing—the champion of classical AI—used dynamical equations (describing chemical diffusion) to define self-organizing systems in which novel structure, such as spots or segmentation, could emerge from a homogeneous origin (see Chapter 5).[12]

Other early members of the movement included the experimental psychologist Kenneth Craik; the mathematician John von Neumann; the neurologists William Grey Walter and William Ross Ashby; the engineer Oliver Selfridge; the psychiatrist and anthropologist Gregory Bateson; and the chemist and psychologist Gordon Pask.[13]

Craik, who died (aged 31) in a cycling accident in 1943, before the advent of digital computers, referred to analogue computing in thinking about the nervous system. He described perception and motor action, and intelligence in general, as guided by feedback from "models" in the brain.[14] His concept of cerebral models, or representations, would later be hugely influential in AI.

Von Neumann had puzzled about self-organization throughout the 1930s, and was hugely excited by McCulloch and Pitts' first paper. Besides changing his basic computer design from decimal to binary, he adapted their ideas to explain biological evolution and reproduction. He defined various cellular automata: systems made of many computational units, whose changes follow simple rules depending on the current state of neighboring units.[15] Some of these could replicate others. He even defined a universal replicator, capable of copying anything—itself included. Replication errors, he said, could lead to evolution.

Cellular automata were specified by von Neumann in abstract informational terms. But they could be embodied in many ways, for example, as self-assembling robots, Turing's chemical diffusion, Beurle's physical waves, or—as soon became clear—DNA.

From the late 1940s on, Ashby developed the Homeostat, an electrochemical model of physiological homeostasis.[16] This intriguing machine could settle into an overall equilibrium state no matter what values were initially assigned to its 100 parameters (allowing almost 400,000 different starting conditions). It illustrated Ashby's theory of dynamical adaptation—both inside the body (not least, the brain) and between the body and its external environment, in trial-and-error learning and adaptive behavior.

Grey Walter, too, was studying adaptive behavior—but in a very different way.[17] He built mini-robots resembling tortoises, whose sensorimotor circuitry modeled Sherrington's theory of neural reflexes. These pioneering situated robots displayed lifelike behaviors such as light-seeking, obstacle-avoidance, and associative learning via conditioned reflexes. They were sufficiently intriguing to be exhibited to the general public at the Festival of Britain in 1951.

Ten years later, Selfridge (grandson of the founder of the London department store) used symbolic methods to implement an essentially parallel-processing system called Pandemonium.[18]

This GOFAI program learned to recognize patterns by having many bottom-level "demons," each always looking out for one simple perceptual input, which passed their results on to higher-level demons. These weighed the features recognized so far for consistency (e.g. only *two* horizontal bars in an **F**), downplaying any features that didn't fit. Confidence levels could vary, and they mattered: the demons that shouted loudest had the greatest effect. Finally, a master-demon chose the most plausible pattern, given the (often conflicting) evidence available. This research soon influenced both connectionism and symbolic AI. (One very recent offshoot is the LIDA model of consciousness: see Chapter 6.)

Bateson had little interest in machines, but he based his 1960s theories of culture, alcoholism, and "double-bind" schizophrenia on ideas about communication (i.e. feedback) picked up earlier at cybernetic meetings. And from the mid-1950s on, Pask—described as "the genius of self-organizing systems" by McCulloch—used cybernetic and symbolic ideas in many different projects. These included interactive theater; intercommunicating musical robots; architecture that learned and adapted to its users' goals; chemically self-organizing concepts; and teaching machines. The latter enabled people to take different routes through a complex knowledge representation, so were suitable for both step-by-step and holistic cognitive styles (and varying tolerance of irrelevance) on the learner's part.

In brief, all the main types of AI were being thought about, and even implemented, by the late 1960s—and in some cases, much earlier than that.

Most of the researchers concerned are widely revered today. But only Turing was a constant specter at the AI feast. For many years, the others were remembered only by some subset of the research community. Grey Walter and Ashby, in particular, were nearly forgotten until the late 1980s, when they were lauded (alongside Turing) as grandfathers of A-Life. Pask had to wait even longer for recognition. To understand why, one must know how the computer modelers became disunited.

How AI Divided

Before the 1960s, there was no clear distinction between people modeling language or logical thinking and people modeling purposive/adaptive motor behavior. Some individuals worked on both. (Donald Mackay even suggested building hybrid computers, combining neural networks with symbolic processing.) And all were mutually sympathetic. Researchers studying physiological self-regulation saw themselves as engaged in the same overall enterprise as their psychologically oriented colleagues. They all attended the same meetings: the interdisciplinary Macy seminars in the USA (chaired by McCulloch from 1946 to 1951), and London's seminal conference on "The Mechanization of Thought Processes" (organized by Uttley in 1958).[19]

From about 1960, however, an intellectual schism developed. Broadly speaking, those interested in *life* stayed in cybernetics, and those interested in *mind* turned to symbolic computing. The network enthusiasts were interested in both brain and mind, of course. But they studied associative learning in general, not specific semantic content or reasoning, so fell within cybernetics rather than symbolic AI. Unfortunately, there was scant mutual respect between these increasingly separate sub-groups.

The emergence of distinct sociological coteries was inevitable. For the theoretical questions being asked—biological (of varying kinds) and psychological (also of varying kinds)—were different. So too were the technical skills involved: broadly defined, logic versus differential equations. Growing specialization made communication increasingly difficult, and largely unprofitable. Highly eclectic conferences became a thing of the past.

Even so, the division needn't have been so ill-tempered. The bad feeling on the cybernetic/connectionist side began as a mixture of professional jealousy and righteous indignation. These were prompted by the huge initial success of symbolic computing, by the journalistic interest attending the provocative term "artificial intelligence" (coined by John McCarthy in 1956 to name what had previously been called "computer simulation"), and by the arrogance—and unrealistic hype—expressed by some of the symbolists.

Members of the symbolist camp were initially less hostile, because they saw themselves as winning the AI competition. Indeed, they largely ignored the early network research, even though some of their leaders (Marvin Minsky, for instance) had started out in that area.

In 1958, however, an ambitious theory of neurodynamics—defining parallel-processing systems capable of self-organized learning from a random base (and error-tolerant to boot)—was presented by Frank Rosenblatt and partially implemented in his photoelectric Perceptron machine.[20] Unlike Pandemonium, this didn't need the input patterns to be pre-analyzed by the programmer. This novel form of connectionism couldn't be ignored by the symbolists. But it was soon contemptuously dismissed. As explained in Chapter 4, Minsky (with Seymour Papert) launched a stinging critique in the 1960s claiming that perceptrons are incapable of computing some basic things.[21]

Funding for neural-network research dried up accordingly. This outcome, deliberately intended by the two attackers, deepened the antagonisms within AI.

To the general public, it now seemed that classical AI was the only game in town. Admittedly, Grey Walter's tortoises had received great acclaim in the Festival of Britain. Rosenblatt's Perceptron was hyped by the press in the late 1950s, as was Bernard Widrow's pattern-learning Adaline (based on signal-processing). But the symbolists' critique killed that interest stone dead. It was symbolic AI which dominated the media in the 1960s and 1970s (and which influenced the philosophy of mind as well).

That situation didn't last. Neural networks—as "PDP systems" (doing parallel distributed processing)—burst onto the public stage again in 1986 (see Chapter 4). Most outsiders—and some insiders, who should have known better—thought of this approach as utterly *new*. It seduced the graduate students, and attracted enormous journalistic (and philosophical) attention. Now, it was the symbolic AI people whose noses were put out of joint. PDP was in fashion, and classical AI was widely said to have failed.

As for the other cyberneticians, they finally came in from the cold with the naming of A-Life in 1987. The journalists, and the graduate students, followed. Symbolic AI was challenged yet again.

In the twenty-first century, however, it has become clear that different questions require different types of answers—horses for courses. Although traces of the old animosities remain, there's now room for respect, and even cooperation, between different approaches. For instance, "deep learning" is sometimes used in powerful systems combining symbolic logic with multilayer probabilistic networks; and other hybrid approaches include ambitious models of consciousness (see Chapter 6).

Given the rich variety of virtual machines that constitute the human mind, one shouldn't be too surprised.