The Mechanical Mind in History

edited by Philip Husbands, Owen Holland, and Michael Wheeler

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

For information about special quantity discounts, please email special_sales@mitpress.mit.edu

This book was set in Stone Serif and Stone Sans on 3B2 by Asco Typesetters, Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

The mechanical mind in history / edited by Philip Husbands, Owen Holland, and Michael Wheeler.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-262-08377-5 (hardcover : alk. paper) 1. Artificial intelligence—History.
2. Artificial intelligence—Philosophy. I. Husbands, Phil. II. Holland, Owen.
III. Wheeler, Michael, 1960–
Q335.M3956 2008
006.309—dc22 2007035271

10 9 8 7 6 5 4 3 2 1

1 Introduction: The Mechanical Mind

Philip Husbands, Michael Wheeler, and Owen Holland

Through myths, literature, and popular science, the idea of intelligent machines has become part of our public consciousness. But what of the actual science of machine intelligence? How did it start? What were the aims, influences, ideas, and arguments that swirled around the intellectual environment inhabited by the early pioneers? And how did the principles and debates that shaped that founding period persist and evolve in subsequent research? As soon as one delves into these questions, one finds oneself enmeshed in the often obscured roots of ideas currently central to artificial intelligence, artificial life, cognitive science, and neuroscience. Here one confronts a rich network of forgotten historical contributions and shifting cross-disciplinary interactions in which various new questions emerge, questions such as: What intellectual importance should we give to little-known corners of the history of the mechanical mind, such as cybernetic art, the frequently overlooked British cybernetic and pre-cybernetic thinkers, and cybernetic influences in politics? And, more generally, how is our understanding of the science of machine intelligence enriched once we come to appreciate the important reciprocal relationships such work has enjoyed, and continues to enjoy, with a broad range of disciplines? Moreover, issues that we sometimes address from within an essentially ahistorical frame of reference take on a new, historicized form. Thus one wonders not "What is the relationship between the science of intelligent machines and the sciences of neuroscience and biology?" but, rather, "In different phases of its history, how has the science of intelligent machines interacted with the sciences of neuroscience and biology?" Of course, once one has taken proper account of the past, the present inevitably looks different. So, having forged a path through the history of the mechanical mind, one is driven to ask: How far have we really come in the search for the mechanization of mind? What have we actually learned? And where should we go next?

The issues raised in the previous paragraph were what inspired, and subsequently drove the development of, the present volume. Unsurprisingly, given the nature and scope of these issues, the volume is essentially and massively cross-disciplinary in character, bringing together papers by scientists, artists, historians, and philosophers. Moreover, some of the best sources of engaging and illuminating insights into any field of study are the personal memories of those who shaped that field. It is here that the drama of science becomes manifest, along with previously undetected connections and influences. To capture these dimensions of our topic, we have chosen to supplement the usual diet of papers with a number of interviews with highly influential thinkers, most of whom were deeply involved in the birth of the field and have been major contributors to it ever since.

So is the mechanization of mind possible? In a sense this is our question, but that sense needs to be carefully specified. We are not focusing here on something analogous to the now-standard distinction between strong and weak artificial intelligence, so our question is not, "Is it possible to build a machine that *really* instantiates mental states and processes as opposed to 'merely' simulating them?" We are interested in the attempt to explain mind scientifically as a wholly mechanical process—mind as, or perhaps as generated by, an intelligent machine. Given that simulations are established weapons in the scientist's explanatory tool kit—in physics, biology, economics and elsewhere—we take this latter issue to be orthogonal to the "real mind versus simulated mind" debate. Second, we are not focusing, at least not principally, on the attempt to mechanize mind in the sense of building a complete functioning mechanical mind, presumably as an aspect of an integrated mobile robotic platform. The primary issue is not the mechanization of *a* mind. Rather, given science's strategy of abstracting to the key elements of a phenomenon in order to explain it, mechanical models of subsets of mind (for instance, mechanical models of individual psychological capacities such as reasoning or perception) are at the heart of the mechanization of mind, in the specific sense of the attempt to explain mind scientifically as a wholly mechanical process. These are the mechanisms that explain mind as machine.

So far, so good. But what sort of machine do we need for this task? This is where things get most interesting, and where, we believe, the present collection makes a genuine intellectual contribution that goes beyond that of historical scholarship. For what the various papers and memoirs here do is illustrate anew the rich kaleidoscope of diverse and interacting notions of mechanism that historically have figured in the shifting landscape of the mechanical mind. In the pages ahead we shall see mind mechanized as an analogue electrical system of wires, valves, and resistors; as an organized suite of chemical interactions; as a self-organizing electromechanical device, as a team of special-purpose mechanisms; as an automated general-purpose information processor; as an abstract deterministic process specified by state-transition rules (such as a Turing machine); as an integrated collection of symbol-manipulating mechanisms; and as an autonomous network of subsymbolic or nonsymbolic mechanisms. We shall see some of these notions deployed in combination as different aspects of the mental machine, and we shall see some of them pitted against each other in debates over the fundamental character of that machine. In addition, we shall see how some of these different notions have influenced and been influenced by the matrix of cross-disciplinary connections identified earlier.

In the remainder of this chapter, the contributions to this book are put into the wider context of the history of mind as machine. This is not intended to be a comprehensive history, or anything like it, but is merely a sketch that helps to show how the chapters relate to each other and to the central themes of the book. This volume offers a wide range of original material, with some emphasis on underexplored areas, such as British cybernetics, and the relationship between the mechanical mind and the arts. It is intended to complement more specific histories (such as those of the cybernetic period, including Heims 1991; Dupuy 2000) as well as more general surveys of the field (McCorduck 1979; Dyson 1997; Cordeschi 2002; and Boden's recent heroic two-volume history of cognitive science [2006]).

Looking at some discussions of the history of artificial intelligence, one would be forgiven for thinking that the mechanization of mind began, or at least took off properly, with the advent of the digital computer and the pioneering work of thinkers such as Allen Newell and Herbert Simon in the second half of the 1950s. But that is a very narrow and ultimately misleading view of history. There is a prehistory of what we now commonly think of as artificial intelligence in the cybernetic movements of the 1940s and 1950s—movements of which Newell and Simon themselves were deeply aware, incidentally. Moreover, there is a pre-prehistory of artificial intelligence that one might reasonably suggest began with (and this will come as a surprise to some readers) René Descartes (1596–1650). Descartes is often portrayed as the archenemy of mind as machine, but in fact he used clocks (relative rarities in his time) and the complex, animal-like automata that (among other things) moved, growled, spoke, and sang for the entertainment of the wealthy elite of seventeenth-century Europe as models for a range of what we would now think of as psychological capacities. Crucially, however, Descartes thought that some psychological capacities, in particular, reason, remained beyond the reach of a "mere" mechanism (Descartes 1637).

Soon afterward, however, the British philosopher Thomas Hobbes (1588–1679) went further than Descartes to become perhaps the first real champion of the mechanization of mind. He played a crucial role in establishing the intellectual climate that would result in attempts to understand the physical processes underlying intelligent behavior, and would later allow the emergence of the modern science of machine intelligence. Although today he is usually remembered as an ethical and political philosopher, Hobbes was one of the most important natural philosophers of his day. His materialist stance emphasized the machinelike qualities of nature, suggesting the possible creation of artificial animals: artificial intelligences and artificial life. In attacking Descartes's separation of mind and body, Hobbes argued that all of human intelligence is the product of physical mechanisms: that mind is a property of suitably organized matter.

The idea of mind as machine, then, stretches back over several centuries. As hinted at above, Descartes was not as hostile to the idea of mechanistic explanations of intelligent behavior as he is often portrayed today. Michael Wheeler explores this theme in some depth in his chapter, "God's Machines: Descartes on the Mechanization of Mind." He shows that Descartes's position was that machines (in the sense relevant to the mechanization of mind) are essentially collections of special-purpose mechanisms, and that no single machine could incorporate the enormous number of special-purpose mechanisms that would be required for it to reproduce human-like behaviour. By looking at contemporary work in biologically-inspired AI, Wheeler asks to what extent we can yet answer Descartes.

Although Hobbes's *Leviathan* included a combinatorial theory of thinking (Hobbes 1651), details of possible mechanisms for intelligence were very sketchy. It was some time before much progress was made in this direction: the eighteenth century saw the construction of many ingenious mechanical automata, including chess-playing Turks and flatulent ducks, but it wasn't until the nineteenth century that major breakthroughs occurred, including the design of Charles Babbage's programmable Analytical Engine.

The son of a London banker, Babbage (1791–1871) was a brilliant mathematician and engineer who held the same chair at Cambridge University that Newton had occupied. Inspired by Leibniz, whose work was in turn influenced by Hobbes, in 1821 he designed his mechanical Difference Engine for calculating accurate mathematical tables—something of enormous practical importance at the time. However, Babbage's interest in calculating machines ran deeper than the production of mathematical tables. He envisioned such engines as powerful tools for science, hoping that their whirring cogs would shed new light on the workings of nature. In this spirit, in 1834 he began work on his revolutionary Analytical Engine, a general, programmable machine. The engine was to read instructions from sets of punched cards, adapted from those used in Jacquard looms (invented in 1801 to automate textile weaving), and to manipulate partial results in its own internal memory. Rather than being designed to perform just one set of calculations, the machine was intended to be a completely general computing engine; in theory, it could be programmed to perform any calculation. In chapter 2, "Charles Babbage and the Emergence of Automated Reason," Seth Bullock explores the context in which Babbage's work emerged, highlighting the debates on the possibility of automated reason, which covered economic, social, and moral ground. He also shows how Babbage was able to demonstrate the wider applicability of his machines by developing the first computational model intended to help further study of a scientific problem (in this case one in geology).

In 1843 Augusta Ada, Countess of Lovelace (1815–1852) translated into English a paper on the Analytical Engine written by the mathematician Luigi Menabrea (Lovelace 1843). Ada was the daughter of Lord Byron, the great poet. Her parents separated almost immediately after her birth, and Lady Byron raised Ada to appreciate mathematics and science, in part because of her own interest in these areas, but also because she hoped it would drive out any Byronic madness her daughter might have inherited. In collaboration with Babbage, Ada added extensive notes to the manuscript, which make it clear that they both understood the importance of the general nature of the Engine. Ada wrote of its potential to act as a "thinking, reasoning machine." The notes include a detailed description of a method for using the Engine to calculate Bernoulli numbers. This is widely regarded as the first computer program, although there is some controversy over whether the primary author was Lovelace or Babbage. Ada was perhaps the first person to see the possibility of using computational engines in the arts, writing of the Analytic Engine's potential to compose music and generate graphics.

The Analytical Engine was never completed; its construction became mired in manufacturing and bureaucratic difficulties that resulted in the British government's withdrawing funding. In 1991 a team at the Science Museum in London constructed the Difference Engine Number 2 according to Babbage's detailed designs. It worked perfectly. In most respects Babbage's remarkable vision of a universal machine anticipated the modern digital computer age by more than a century.

While Babbage was struggling to construct his engines, the English mathematician George Boole (1815–1864), the self-educated son of a Lincoln cobbler, was building a formal system of logic which went on to serve as a cornerstone of all modern digital technology, but which was also intended to capture the structure of reasoning and thinking (Boole 1854). In Boolean algebra, logical relationships between entities are formalized and manipulated. Variables representing the entities are restricted to two possible values, true or false—1 or 0. By uniting logic with mathematics, in particular binary arithmetic, Boole laid the foundations for the flow of bits and bytes that power our digital age. He died after developing a fever following a soaking in a rainstorm. His demise was unwittingly aided by his wife, who, believing that a cure should mirror the cause, threw buckets of cold water over him as he lay shivering in bed.

Where Babbage and his predecessors developed schemes for describing and automating reasoning at a fairly high, abstract level, one of the first people to try to ground intelligence in brain function was Alfred Smee (1818–1877), a brilliant scientist and engineer who held the somewhat bizarre position of surgeon to the Bank of England. (His father was secretary of the bank and the position was specially created in the hope of tapping into Alfred's inventive flair. It did: he developed electrotype plate printing of banknotes, which greatly reduced problems with forged notes.) Smee pioneered theories of the operation of the nervous system, speculating on how its electrical networks were organized. He also formulated ideas about artificial sense organs and a type of very early artificial neural network.

During the early decades of the twentieth century, advances in electrical engineering and early electronics fed into formal theories of the operation of neurons, as well as greatly improving experimental techniques in the developing field of neurophysiology. This allowed great pioneers such as Lord Adrian (1889–1977) and Charles Sherrington (1857–1952) to lay the foundations for the modern view of the nervous system by greatly advancing knowledge of the electrical properties of nerve cells (Adrian 1928; Sherrington 1940). Communications theory was also emerging in engineering circles; as we shall see, future developments in this area would later have a significant impact on approaches to the mechanization of mind.

At about the same time that Adrian and Sherrington were making great strides in understanding neurons, D'Arcy Thompson was trying to fathom how biological structures develop. In 1917 he published his celebrated book

On Growth and Form (Thompson 1917). As Margaret A. Boden argues in chapter 3, "D'Arcy Thompson: A Grandfather of A-Life," this pioneering work of mathematical biology, in which Thompson sought to develop a quantitative approach to biological forms and processes of growth, not only helped to pave the way for modern theoretical biology but also prefigured the contemporary field of artificial life (or A-Life), the study of life in general, abstract terms. As well as influencing Alan Turing's work on morphogenesis, of which more later, it emphasized the embodied nature of natural intelligence, a theme that has become increasingly central to contemporary cognitive science (Pfeifer and Scheier 1999; Wheeler 2005).

The notion of embodied mechanical intelligence was, quite literally, thrust center stage in the years between the world wars, when Karel Čapek's play *R.U.R.* introduced the world to robots, in the process forging the associated myths and images that now permeate our culture. In "The Robot Story: Why Robots Were Born and How They Grew Up," Jana Horáková and Jozef Kelemen give a detailed account of the origins of Čapek's work, tracing its roots to the dreams and folk tales of old Europe. They show how it was a product of its troubled times and how the idea of robots was interpreted in different ways in Europe and America as it seeped into the collective unconscious. The new dreams and images thus created undoubtedly inspired future generations of machine intelligence researchers.

Smee's early desire to unite the workings of the mind with the underlying neural mechanisms, and to develop machines around the principles uncovered, was a theme that reemerged very strongly in the mid-twentieth century. It was in this period that machine intelligence really took off. At the same time advances in understanding the nervous system continued apace. Kenneth Craik (1914–1945) was an influential, if now often forgotten, figure in the flurry of progress that occurred. Craik was a brilliant Scottish psychologist, based at Cambridge University, who pioneered the study of human-machine interfaces, and was a founder of cognitive psychology and also of cybernetic thinking. He died tragically young, in a road accident on the last day of the war in Europe, his potential surely not fully realized. His classic 1943 book, *The Nature of Explanation* (Craik 1943), introduced the radical and influential thesis that the brain is a kind of machine that constructs small-scale models of reality that allow anticipation of external events. Disgruntled with mainstream philosophy of mind and much of psychology, and inspired by the strides Adrian and his colleagues were making, he maintained that explanations of intelligence should incorporate an understanding of the underlying neural processes. Craik's influence on the development of cybernetics, on both sides of the Atlantic, is discussed in Philip Husbands and Owen Holland's chapter on the Ratio Club.

At the same time as Craik was starting to develop his ideas, in another part of Cambridge the mathematician Alan Turing (1912–1954) was about to publish a startling paper on one of David Hilbert's open problems in mathematics, the *Entscheidungsproblem* (''decision problem''), namely: Is it possible to define a formal procedure that could be used to decide whether any given mathematical assertion was provable. Turing's highly original approach to the problem was to define a kind of simple abstract machine (Turing 1936). By using such a machine as a very general way of constructing a formal procedure in mathematics, he was able to show that it followed that the answer to the problem was no. The concept of the Turing machine, as it became known, now serves as the foundation of modern theories of computation and computability. In the paper Turing explicitly drew a parallel between the operation of such a machine and human thought processes. Turing also introduced a more general concept that was to have an immense practical impact: the Universal Turing Machine. This machine could interpret and then execute the set of instructions defining *any* given standard Turing machine (each of which corresponded to a particular formal procedure or algorithm). Thus, the Universal Turing Machine embodies the central principle of the computer as we know it today: a single machine that can perform any well-defined task as long as it is given the appropriate set of instructions, or program. A hundred years after Babbage, and by a very different route, Turing envisaged a completely general supermachine. This time the vision was to come to fruition.

Donald Michie's chapter, ''Alan Turing's Mind Machines,'' draws on his experience as one of Turing's close colleagues in wartime code-cracking work at Bletchley Park, the headquarters of Britain's cryptography efforts, to give insights into the development of Turing's ideas and the early computers that flowed from them. He argues that Turing's unfashionable and often resisted obsession with tackling combinatorial problems with brute-force computation, partly born of his wartime experience with cryptanalytical problems, helped to shape the way computers came to be used. He shows that computer analyses of combinatorial domains such as chess, inspired by Turing's work, are still of great importance today in yielding new approaches to the difficult problem of transparency in complex computer-based decision systems.

In a complementary chapter, Andrew Hodges asks ''What did Alan Turing Mean by 'Machine'?'' He focuses on the title of Turing's unpublished 1948 report "Intelligent Machinery" (Turing 1948) to explore what Turing intended by an "intelligent machine." Turing saw central roles for the new digital computers in the development of machine intelligence and in the exploration of brain mechanisms through simulations, both of which came to pass. Hodges argues that although the central thrust of Turing's thought was that the action of brains, like that of any machine, could be captured by classical computation, he was aware that there were potential problems in connecting computability with physical reality.

The Second World War was to prove a major catalyst for further advances in mechanistic conceptions of intelligence as well as in the development of practical computers. In Britain there was little explicitly biological research carried out as part of the war effort, so most biologists were drafted into the main thrust of scientific research on communications and radar. As explained in chapter 6, this was to have the extremely important effect of exposing these biologists to some electronics and communication theory as well as to engineers and mathematicians who were experts in these areas. This mixing of people and disciplines led to an important two-way flow of ideas that was to prove highly significant in advancing the formal understanding of the nervous system as well as developments in machine intelligence. There was much discussion of electronic brains, and the intense interest in the subject carried over into peacetime.

In the early 1940s a circle of scientists intent on understanding general principles underlying behavior in animals and machines began to gather around the MIT mathematician Norbert Wiener (1894–1964). Inspired by Wiener's classified work on automatic gun aiming, Arturo Rosenblueth, Wiener, and Julian Bigelow (1943) published a paper on the role of feedback mechanisms in controlling behavior. This work triggered great interest among other American scientists in new approaches to the mechanization of mind. Influenced by Wiener's ideas, but also aware of Craik's and Turing's work, the group was initially composed of a small number of mathematicians and engineers (Wiener, John von Neumann, Bigelow, Claude Shannon, Walter Pitts) and brain scientists (Rafael Lorente de Nó, Rosenblueth, Warren McCulloch). A series of meetings sponsored by the Macy Foundation saw the group expand to incorporate the social sciences. Wiener named the enterprise cybernetics; the publication of his book *Cybernetics, or Control and Communication in the Animal and the Machine* (Wiener 1948), along with the proceedings of the Macy meetings (von Foerster 1950–55), did much to spread its influence and popularity. As well as Wiener's book, notable developments that came under the cybernetic umbrella included McCulloch and Pitts's seminal work on mathematical descriptions of neuronal networks (McCulloch and Pitts 1943; Pitts and McCulloch 1947), providing the first examples of artificial neural networks, and Shannon's information theory (Shannon and Weaver 1949). McCulloch and Pitts modeled neuronal networks in terms of connected logic units and showed that their nets were equivalent to Universal Turing Machines, implicitly suggesting a close link between the nervous system and the digital computer. Information theory, which provided a mathematical framework for designing and understanding communication channels, is another foundation stone of the digital age. It also provided new ideas about the operating principles of biological senses and what kinds of processing might be going on in the nervous system.

In Britain, where war work had also familiarized many scientists with feedback mechanisms and early information theory, a parallel group formed, the Ratio Club. The club was founded and organized by John Bates, a neurologist at the National Hospital for Nervous Diseases in London. The other twenty carefully selected members were a mixed group of mainly young neurophysiologists, engineers, and mathematicians, with the center of gravity firmly toward the brain sciences. This illustrious group included W. Ross Ashby, Horace Barlow, Thomas Gold, Jack Good, Donald MacKay, Alan Turing, W. Grey Walter, and Albert Uttley. Most members had a strong interest in developing "brainlike" devices, either as a way of formalizing and exploring theories about biological brains, or as a pioneering effort in creating machine intelligence, or both. Most meetings of the club occurred between September 1949 and July 1953. During this extremely productive period various members made highly significant contributions to cybernetics and related fields. Husbands and Holland's chapter, "The Ratio Club: A Hub of British Cybernetics," for the first time tells the story of this remarkable group. Horace Barlow's very significant contributions to neuroscience, including his introduction into it of important information-theoretic concepts (Barlow 1959), were heavily influenced by the club. Members pioneered a wide range of techniques and ideas that are proving to be ever more influential. For instance, Grey Walter (1910–1977), a leader in electroencephalographic (EEG) research, built the first autonomous mobile robots, controlled by simple electronic nervous systems (Walter 1953). W. Ross Ashby (1903–1972), who had actually published on the role of feedback in adaptive systems several years before Rosenblueth, Wiener, and Bigelow (Ashby 1940), further developed such notions, culminating in their demonstration in his adaptive Homeostat machine (Ashby 1952); and Turing, whose seminal paper on machine intelligence (Turing 1950) was published during the club's lifetime, pioneered the use of computational models in biology in his groundbreaking work on morphogenesis, which showed how regular patterns could be formed by appropriately parameterized reaction-diffusion systems—work that called up the spirit of D'Arcy Thompson (Turing 1952).

Ashby, who is now widely acknowledged as the most important theorist of cybernetics after Wiener—partly through the influence of his books (Ashby 1952, 1956)—had a singular vision that he had developed in isolation for many years before becoming part of the scientific establishment in the late 1940s. His unique philosophy, which stressed the dynamic nature of brain mechanisms and the interactions between organism and environment, is explored by Peter Asaro in chapter 7, ''From Mechanisms of Adaptation to Intelligence Amplifiers: The Philosophy of W. Ross Ashby.'' Asaro sheds light on what kind of machine Ashby thought the brain was and how its principles might be captured in an artificial device.

Parallel developments in the United States also focused on biologically inspired brainlike devices, including work by researchers such as Frank Rosenblatt and Marvin Minsky on the construction of electronic artificial neural networks that were able to perform simple learning tasks. Oliver Selfridge, a grandson of the founder of London's famous Selfridge's department store, had left Britain at the age of fourteen to study with Wiener at MIT. In the mid-1950s he developed his breakthrough Pandemonium system, which learned to recognize visual patterns, including alphanumeric characters (Selfridge 1959). The system employed a layered network of processing units that operated in parallel and made use of explicit feature detectors that only responded to certain visual stimuli—a more general mechanism than the specific detectors that had recently been shown to exist in biological vision systems by Horace Barlow in the form of ''fly detectors'' in the frog's retina (Barlow 1953). Neural mechanisms that are selectively responsive to certain general features (for instance, edge and convexity detectors) were subsequently shown to exist in natural vision systems by Jerry Lettvin, Humberto Maturana, Warren McCulloch, and Walter Pitts (1959).

Most prominent among the second wave of British cyberneticists were Stafford Beer (1926–2002) and Gordon Pask (1928–1996), who were both particularly influenced by Ashby. Beer took cybernetic ideas into the world of industrial management and became a highly successful consultant to corporations and governments alike. In ''Santiago Dreaming,'' Andy Beckett tells the story of how in the early 1970s the Allende administration in Chile engaged Beer to design and develop a revolutionary electronic communication system in which voters, workplaces, and the government were to be linked together by a kind of "socialist internet." Pask was an eccentric figure who strode around in an Edwardian cape while pursuing radical ideas far from the mainstream. In "Gordon Pask and His Maverick Machines," Jon Bird and Ezequiel Di Paolo highlight Pask's willingness to explore novel forms of machine, often in collaboration with Beer, in his quest to better understand principles of self-organization that would illuminate the mechanisms of intelligence. These included a "growing" electrochemical device intended to act as an artificial ear. They show how Pask's work is relevant to current research in AI and A-life, and how key questions he posed have not yet been answered.

Pask, like other machine intelligence researchers before and since, was interested in applying his ideas in the visual arts. As Paul Brown shows in chapter 11, "The Mechanization of Art," Wiener's and Ashby's ideas were quickly appreciated by a number of artists, such as Nicolas Schöffer, who in the mid-1950s pioneered a kind of autonomous kinetic art, cybernetic sculptures. Brown traces the cultural, as well as scientific, antecedents of this work in an account of how the mechanization of art developed over the centuries. He focuses on its growth during part of the second half of the twentieth century, a period that saw the influential 1968 Institute of Contemporary Arts (London) exhibition *Cybernetic Serendipity*, which featured Pask's installation *Colloquy of Mobiles*. He reminds us that a number of artists working in this field, such as Edward Ihnatowicz (1926–1988), pioneered approaches to autonomous systems, prefiguring today's growing dialogue between artists and scientists in this area.

In 1956 two young American academics, John McCarthy and Marvin Minsky, organized a long workshop at Dartmouth College to develop new directions in what they termed *artificial intelligence*. McCarthy in particular proposed using newly available digital computers to explore Craik's conception of intelligent machines as using internal models of external reality, emphasizing the power of symbolic manipulation of such models. At the workshop, Allen Newell (1927–1992) and Herbert Simon (1916–2001), influenced by aspects of Selfridge's work, demonstrated a symbolic reasoning program that was able to solve problems in mathematics. This was the beginning of the rise of logic-based, symbol-manipulating computer programs in the study of machine intelligence. This more abstract, software-bound paradigm came to dominate the field and pulled it away from its biologically inspired origins. For a while the term "artificial intelligence," or AI, was exclusively associated with this style of work. This paradigm, which to some extent harked back to the older ideas of Boole and Leibniz, also served as a new kind of abstract model of human reasoning, becoming very influential in psychology and, later, in cognitive science.

Roberto Cordeschi illustrates some of the tension between cybernetic and early AI theories in his chapter, "Steps Toward the Synthetic Method: Symbolic Information Processing and Self-Organizing Systems in Early Artificial Intelligence Modeling." He compares two theories of human cognitive processes, one by the Ratio Club member and cyberneticist Donald Mackay (1922–1987), the other by Newell and Simon. MacKay's model is constructed around his notion of self-organizing systems, whereas Newell and Simon's is based on high-level symbol manipulation. Cordeschi explores epistemological issues raised by each.

The new AI movement in the United States gained significant financial and industrial support in the 1960s, as it began to dominate the arena while the influence and impetus of cybernetics fell away. However, work in neural nets, adaptive and self-organizing systems, and other outgrowths of cybernetics did not disappear altogether. As the weaknesses of the mainstream AI approaches became apparent and the adaptive-systems methods improved, with a number of crucial advances in artificial neural networks and machine learning, the tide turned (see Anderson and Rosenfeld 1998 for an excellent oral history of the rise and fall and rise of artificial neural networks). Since the late 1980s, biologically inspired and subsymbolic approaches have swept back to take center stage. These include an emphasis on whole embodied artificial "creatures" that must adapt to real unforgiving environments. Their brains run on onboard digital computers, as Turing foresaw more than fifty years ago. Work in machine intelligence has again become much more closely aligned with research in the biological sciences. Many of the ideas and methods developed by the great pioneers of the mid-twentieth century have once more come to the fore—the mechanization-of-mind project, although still very far from completion, appears to be back on track. Which is not to say that there is agreement on the best way forward.

One of the most prominent critics of classical AI, or good old-fashioned AI—GOFAI—was Hubert Dreyfus. In "Why Heideggerian AI Failed and How Fixing It Would Require Making It More Heideggerian," he turns the spotlight on one of GOFAI's replacements. Informed by personal experiences and encounters at MIT (the high temple of AI, new and old), Dreyfus tells of how he watched the symbol-processing approach degenerate, and of how it was replaced by what he terms "Heideggerian AI," a movement that began with the work of Rodney Brooks and colleagues (Brooks 1999).

This work puts central emphasis on acting in the world and thus concentrates on the development of mobile autonomous robots. Dreyfus explains why, in his view, this style of AI has also failed and suggests how it should be fixed, calling on Walter Freeman's neurodynamics and stressing the importance of the specifics of how particular bodies interact with their environments.

The final section of the book offers a series of interviews, conducted by one of the editors, with major figures whose careers were firing into life in the middle of the last century, an astonishingly fertile period in the search for the secrets of mechanical intelligence. We are given vivid accounts of how these great scientists' ideas developed and of who influenced them. Certain themes and characters echo through these interviews, giving fresh perspective on material earlier in the book.

John Maynard Smith, one of the great evolutionary biologists of the twentieth century, who originally trained as an engineer, gives us an insight into the spirit of science immediately after the Second World War as well as into the early influence of cybernetics on developmental and evolutionary biology. John Holland, the originator of genetic algorithms, recounts how his theories of adaptive systems were in turn influenced by biology, then reflects on recent developments and considers why, in the late 1980s, there was a great resurgence of interest in complex adaptive systems. Oliver Selfridge, one of the pioneers of machine learning, tells us what it was like to be at the heart of the MIT cybernetics enterprise in the 1940s and 1950s, and how he helped Minsky and McCarthy to establish the field of AI. Regretting GOFAI's lack of interest in learning and adaptation during its heyday, he gives his views on where the field should go now. The great neuroscientist Horace Barlow paints a picture of life in Lord Adrian's department at Cambridge University during the late 1940s and tells how the Ratio Club profoundly influenced his subsequent career. Toward the end of his interview he makes the highly pertinent point that as neuroscience has developed over the past fifty years, it has fragmented into specialized subareas. So although knowledge has increased to an enormous extent, there is now a greater need than ever for an overarching theory. The theorists, experimentalists, and modelers must all combine in a coherent way if we are ever to understand the nervous system in sufficient detail to formulate its principles. Jack Cowan, a pioneer of neural networks and computational neuroscience, gives a unique perspective on activity in machine intelligence in the UK and the United States in the late 1950s and early 1960s. He recounts how his ideas developed under the influence of some of the great pioneers of cybernetics, and how those ideas flourished throughout his subsequent career.

From positions of authority, with access to extraordinarily wide perspectives, these pioneers look back at what has been achieved, and comment on how far we still have to go, in the mechanization of mind. All are optimistic for the long term, but stress the enormous complexity of the task. In short, although much has been achieved and great progress has been made in understanding the details of specific mechanisms and competences, in terms of the overall picture, we have not yet come very far at all. This message serves as a useful antidote to the wild ravings of those who claim that we will soon be downloading our minds into silicon (although it is not clear whether this will be before or after our doors are kicked in by the superintelligent robots that these same people claim will take over the world and enslave us).

References

Adrian, Edgar Douglas (Lord Adrian). 1928. *The Basis of Sensation*. London: Christophers.

Anderson, J. A., and E. Rosenfeld. 1998. *Talking Nets: An Oral History of Neural Networks*. Cambridge, Mass.: MIT Press.

Ashby, W. Ross. 1940. "Adaptiveness and Equilibrium." *Journal of Mental Science* 86: 478.

———. 1952. *Design for a Brain*. London: Chapman & Hall.

———. 1956. *An Introduction to Cybernetics*. London: Chapman & Hall.

Barlow, Horace B. 1953. "Summation and Inhibition in the Frog's Retina." *Journal of Physiology* 119: 69–88.

———. 1959. "Sensory Mechanism, the Reduction of Redundancy, and Intelligence." In *Mechanisation of Thought Processes: Proceedings of a Symposium held at the National Physical Laboratory on 24–27 November 1958*, edited by D. Blake and Albert Uttley. London: Her Majesty's Stationery Office.

Boden, M. A. 2006. *Mind as Machine: A History of Cognitive Science*. Oxford: Oxford University Press.

Boole, G. 1854. *An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities*. London: Macmillan.

Brooks, R. A. 1999. *Cambrian Intelligence: The Early History of the New AI*. Cambridge, Mass.: MIT Press.

Cordeschi, Roberto. 2002. *The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics*. Dordrecht: Kluwer Academic Publishers.

Craik, Kenneth J. W. 1943. *The Nature of Explanation*. Cambridge: Cambridge University Press.

Descartes, René. 1637/1985. "Discourse on the Method of Rightly Conducting One's Reason and Seeking the Truth in the Sciences." In *The Philosophical Writings of Descartes*, edited by J. Cottingham, R. Stoothoff, and D. Murdoch. Volume 1. Cambridge: Cambridge University Press.

Dupuy, J.-P. 2000. *The Mechanization of the Mind*. Translated from French by M. B. DeBevoise. Princeton: Princeton University Press.

Dyson, George. 1997. *Darwin Among the Machines*. Reading, Mass.: Addison-Wesley.

Foerster, H. von, ed. 1950–55. *Cybernetics, Circular Causal and Feedback Mechanisms in Biological and Social Systems: Published Transactions of the 6th, 7th, 8th, 9th and 10th Conferences*. 5 volumes. New York: Josiah Macy Jr. Foundation.

Heims, S. 1991. *Constructing a Social Science for Postwar America: The Cybernetics Group, 1946–1953*. Cambridge, Mass.: MIT Press.

Hobbes, Thomas. 1651. *Leviathan*, London: Andrew Crooke.

Lettvin, Jerry Y., H. R. Maturana, Warren S. McCulloch, and Walter H. Pitts. 1959. "What the Frog's Eye Tells the Frog's Brain." *Proceedings of the IRE* 47: 1940–59.

Lovelace, Ada. 1843. "Notes on L. Menabrea's 'Sketch of the Analytical Engine Invented by Charles Babbage, Esq.'" In *Taylor's Scientific Memoirs*. Volume 3. London: J. E. & R. Taylor.

McCorduck, P. 1979. *Machines Who Think: A Personal Inquiry into the History and Prospect of Artificial Intelligence*. San Francisco: Freeman.

McCulloch, Warren S., and Walter Pitts. 1943. "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics* 5: 115–33.

Pfeifer, R., and C. Scheier. 1999. *Understanding Intelligence*. Cambridge, Mass.: MIT Press.

Pitts, Walter, and Warren S. McCulloch. 1947. "How We Know Universals: The perception of Auditory and Visual Forms." *Bulletin of Mathematical Biophysics* 9: 127–47.

Rosenblueth, A., Norbert Wiener, and J. Bigelow. 1943. "Behavior, Purpose and Teleology." *Philosophy of Science* 10, no. 1: 18–24.

Selfridge, Oliver G. 1959. "Pandemonium: A Paradigm for Learning." In *The Mechanisation of Thought Processes*, edited by D. Blake and A. Uttley. Volume 10, *National Physical Laboratory Symposia*. London: Her Majesty's Stationery Office.

Shannon, Claude, and W. Weaver. 1949. *The Mathematical Theory of Communication*. Chicago: University of Illinois Press.

Sherrington, Charles. 1940. *Man on His Nature*. Oxford: Oxford University Press.

Thompson, D. W. 1917. *On Growth and Form*. Cambridge: Cambridge University Press.

Turing, Alan M. 1936. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42, no. 2: 230–65.

———. 1948/2004. "Intelligent Machinery." Report for the National Physical Laboratory. In *The Essential Turing*, edited by A. M. Turing and B. J. Copeland. Oxford: Oxford University Press. Available at www.turingarchive.org.

———. 1950. "Computing Machinery and Intelligence." *Mind* 59: 433–60.

———. 1952. "The Chemical Basis of Morphogenesis." *Philosophical Transactions of the Royal Society of London* (series B) 237: 37–72.

Walter, W. Grey. 1953. *The Living Brain*. London: Duckworth.

Wheeler, Michael. 2005. *Reconstructing the Cognitive World*. Cambridge, Mass.: MIT Press.

Wiener, Norbert. 1948. *Cybernetics, or Control and Communication in the Animal and the Machine*. Cambridge, Mass.: MIT Press.