Minds and Machines (2005) 00:1–5
DOI 10.1007/s11023-005-2624-6

Short Communication

Book Review

Roberto Cordeschi, *The Discovery of the Artificial: Behaviour, Mind and Machines Before and Beyond Cybernetics*, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002, xx + 312, ISBN 1-4020-0606-3.

Italian philosopher Roberto Cordeschi has been writing on the history of artificial intelligence and cybernetics for over a decade. His book *The Discovery of the Artificial: Behaviour, Mind and Machines Before and Beyond Cybernetics*, published in 2002 by Kluwer, provides a thorough, 312 pages long, monograph into the history of the concept of the artificial. While the history of artificial intelligence has been described relatively often (most recently in Crevier (1993), Chrisley (2000), and McCorduck (2004)), *The Discovery of the Artificial* provides a rather unique focus on the pre-cybernetic period, roughly covering the first half of the previous century, before the "birth" of AI at the 1956 Dartmouth conference. When reading about artificial intelligence or its historical background, this pre-cybernetic period is often ignored. The reason for the neglect of this period may well be the lack of notable "famous" achievement from this period, like Turing's test and machine, or Newell and Simon's General Problem Solver. Indeed, it is not the actual inventions of the cyberneticians that speak to our imagination; it is the way they talked about them. In particular the first four chapters of Cordeschi's book illuminate how the heated debates, the formulation of goals and the discussions about terminology paved the road for, and are sometimes very similar to, current discussions in AI, philosophy and cognitive science. The combination of the clear writings about this discourse, and the detailed account of technical issues make this book enjoyable for a wide range of scholars, from the humanities and social sciences, who can learn to appreciate the basics of our current thinking about cognition, to computer scientists, AI practitioners, and philosophers, whose outlook is usually directed at the future, but who will be surprised at the ancientness of their concepts, and the nature of the machines that once were the object of those very same concepts.

The book starts at the beginning of the last century, and chronologically speaking does not go back much further. This may come as a surprise, since "the artificial" can clearly be discovered in the work of Babbage, or even of philosophers like Descartes, Hobbes and Leibniz. Cordeschi defends this choice by pointing out that this history has already been told several times,

and that the artificial as understood in the present book, namely a new concept of a machine that aims to overcome the traditional oppositions between inorganic and organic worlds, between laws that govern the behaviour of physical systems and those that govern the behaviour of organisms, and between causal and teleological explanation, is remote from the automata of the 18th century. True as that may be, the ''ancient'' history of AI, much like the current field of AI, did not only revolve around manufacturing actual machines, but also introduced ideas that influenced contemporary conceptions of the artificial (see Chrisley (2000)).

Still, the roots of assumptions of modern artificial intelligence stemming from the beginning of this century, the so-called ''culture of the artificial'', are dealt with very thoroughly. Cordeschi aims to identify turning points that mark the discovery of the artificial, and to present these in their actual context. This will not only make them more comprehensible, but also avoids slipping into an anachronistic discourse. Incidentally, this also makes the book very vivid and readable, since abstract concepts are illuminated with tangible and rare designs of early ''thinking'' machines. Chapter one starts with a detailed account of disputes on reductionism versus vitalism, intelligence versus automatism, and the use of mentalistic language, mainly held between Loeb and Jennings. However, these disputes did not involve machines. They regarded animal behaviour. The subsequent change of the discourse from animal behaviour to inorganic machines is illuminated by detailed explanations and sketches, for instance of Loeb's phototropic robot, or Hammond's electric dog, which, at the time, was awarded almost superhuman intelligence. The arguments used in these discussions were later echoed in behaviourist psychology. The chapter further shows the confusion on the status of arguments and aims of the field, and the conceptual development of terms like animal, machine or intelligence. These topics are also linked to current thought. For instance, the goals and usefulness of mentalistic language, another important issue raised at the time, are linked to Dennett's intentional stance. The topics seem to be raised accidentally, but they are followed up on throughout the remainder of the book. However, as in the next chapter, in the midst of this explosion of ideas, the reader would sometimes benefit from a more abstract level of analysis and structure of the content.

Chapter two, one of the most elaborated chapters, proceeds with an account of nervous conduction, introducing neurological hypotheses on associationism and connectionism. Again the ancientness of the dispute, which includes early attempts to test hypotheses about the nature of memory, adaptation and learning through mechanical models, is illuminating to all those interested in cognitive science. An outline is given of the link between neurological hypotheses and machines, and it is interesting to see how the nature of the connection between mental associations and nervous connec-

tions was dealt with at the time (according to some this took place at the synapse endings). The mix of languages from newborn neurophysiology and traditional psychological idiom resulted in a hydraulic analogy for the nervous system. The usefulness of analogies within the incomplete state of neurophysiologic knowledge at the time was appreciated by associationist Meyer, who in fact introduced the drainage analogy. He saw mechanical analogies as teaching aids, which help to understand quantitative properties of neurons without having to invoke "ghostlike" entities. The specification of these analogies and the vivid accounts of ancient attempts to penetrate the concept of nervous conduction are of great historical value. The focus on these ancient, forgotten metaphors that rely heavily on the technological advances of their time makes one wonder how we will view today's metaphors a 100 years from now.

Furthermore, this chapter includes early views on connectionism, involving the general idea that when an inorganic machine behaves as predicted by the behavioural theory whose essential elements it embodies, this argues in favour of that theory. It is interesting to see what role was imputed to the machine by different people. Where psychologist McDougal saw mechanism only as automaton theory, biologist Von Euxküll saw abstract machines as fictitious schemata, and compared them to a sign language by means of which it is possible to express new results in a clear way. Engineer Russell was aware that his machines could only simulate simple kinds of learning, but also said that the existence of those machines offered an argument for the sufficiency of the physical principles involved. Like the models discussed in chapter one, these connectionist models were based on the idea of a machine capable of reorganizing its own internal structure. They differed from views such as those of Loeb because they conceived of learning as a particular form of automatism that even inorganic machines could achieve, thereby bridging the gap between the organic and the inorganic. In fact, in 1913 Russell predicted a new discipline of engineering and psychology and physiology, working together at deciphering the nervous system underlying intelligent behaviour.

Chapter three, entitled "The Robot Approach", describes Hull's synthetic approach in the 1920s and 30s that involved the simulation of relations between essential functions. He attempted to verify behaviourism by building machines that simulate simple learning processes, such as Pavlov's conditioned reflex learning. Ultimately, he aimed to manufacture psychic machines. This was the first time that a coherent research programme with an interdisciplinary calling was formed, raising the hope of progressively building intelligent machines. Hull initiated the first attempts to make a detailed comparison between the performance of an organism and an AI device, influencing, among others, biophysicist Rashevsky, who predicted in 1931 that machines might talk, learn and even be intentional. Ross, an

engineer famous for his automata, based his designs on Hull's ideas, and stated that the behaviour of his robots was truly intelligent. These robots could learn mazes (but never failed twice, which was not very much like human intelligence). Ross, in turn, was followed by well known robot designers like Walter and Ashby, but also some less familiar followers of Hull's approach are discussed. Later, Hull turned his back on psychic machines, probably due to the scientific climate that was unfriendly towards his ideas.

Borrowing its title from the famous paper of Rosenblueth et al. (1943), usually regarded as pivotal to cybernetics, chapter four describes the attempt to reinstate teleology within scientific knowledge. Some little known aspects of the discussion in science and philosophy about teleology are presented. Craik is discussed as marking the transition from the mechanistic stance of behavioural sciences in the 1920s to the development of theory and technology of automatic control in the 1940s. Where an intelligent automaton was once termed a contradiction in terms, some artefacts now suggested automatism and control in living organisms and in machines.

Previous chapters dealt with the various stages in the discovery of a novel strategy in the study of mind and behaviour, suggested by new notions of the machine. Chapters five and six explore more recent and better known developments in artificial intelligence, discussing themes that can be traced back to the pre-cybernetic period. A wide range of topics is addressed in this chapter, varying from cybernetic robotics and information processing psychology to old and new connectionism and artificial life. The strong points of this chapter are the connections made between "new" artificial intelligence and the notions of the artificial that were manifest in the pre-cybernetic period. This issue could have been emphasised even more, to the expense of repeating classic AI history.

The last chapter provides a useful structure for discussing the various topics that have been touched upon. Five theses are identified that have been apparent throughout the book: functionalism, modelling method, representationalism, mentalism, and identity of explanatory principles. In his 1991 publication on the same topic, Cordeschi presented this structure directly. Though the story told in the present book unfolds itself logically and sometimes reads like a novel, an outline of the above mentioned theses in the first chapter would have been helpful to put the many ideas into some perspective, in particular to those who are not familiar with recurrent themes in early or contemporary artificial intelligence. In the introduction Cordeschi states that he aims to tell three stories: of strong AI, of weak AI and of the implications of epistemological questions on the mind-body problem, in other words, the theorizing about cognition. This aim was high. However, Cordeschi has been able to come up with a book that will help those studying social sciences to realize the mechanical roots of the ongoing debate about

cognition, but at the same time will help scholars of AI to put their discourse on intelligence in a historical and social perspective.

References

Chisley R. (ed.) (2000), *Artificial Intelligence: Critical Concepts*, London: Routledge.

Cordeschi, R. (1991), 'The Discovery of the Artificial: Some Protocybernetic Developments 1930—1940', *AI & Society* 5, pp. 218–338.

Crevier, D. (1993), *AI: The Tumultuous History of the Search for Artificial Intelligence*, New York: Basic Books of Harper Collins Publishers.

McCorduck, P. (2004), *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence*, Natick, MA: A.K. Peters.

Rosenblueth, A., Wiener, N., & Bigelow, J. (1943), 'Behaviour, purpose and teleology', *Philosophy of Science* 10, pp. 18–24.

SANDER BEGEER
Developmental Psychology
Vrije Universiteit-Amsterdam
Van der Boechorststraat 1, 1081, BT, Amsterdam
The Netherlands
E-mail: S.Begeer@psy.vu.nl