Climbing Mount AI

The Deep Learning Revolution

By Terrence J. Sejnowski

MIT PRESS: 2018.
352 PP. £24.00

Terrence Sejnowski is justifiably known as one of the fathers of artificial neural networks, the computational engines that power modern artificial intelligence (AI) systems. As director of the Computational Neurobiology Laboratory at the Salk Institute, and founder of the highly respected journal *Neural Computation*, Sejnowski has laboured hard to support research into artificial neural networks. In the past ten years, that research has borne substantial amounts of fruit. The field of artificial neural networks, once the bastion of a handful of researchers, has burgeoned into deep learning neural networks. According to Sejnowski, we have already begun to build a brave new world where deep learning AI systems in smartphones are used to recognize speech, where translating between spoken languages is achieved in real time, and where driverless cars will soon become commonplace.

This book is partly a memoir, but is primarily the story of Sejnowski's long trek to the summit of Mount AI. Sejnowski has met many of the leading scientists of the late twentieth century, and has published innovative research papers with a few of them. Photographs of Sejnowski's fellow trekkers put faces to names that are famous within (and occasionally outside of) the AI community. Readers interested in how scientific theories develop will be intrigued to hear how characters like Francis Crick (co-discoverer of DNA) influenced Sejnowski's ideas. Sejnowski's long collaboration with Geoff Hinton comes across strongly, and their early achievements are probably more than either of them could have accomplished alone; they are the Lennon and McCartney of neural networks. Of course, others have played a major part in establishing neural networks as a force to be reckoned with, but Hinton and Sejnowski deserve a substantial proportion of the credit.

Sejnowski tells of the resistance, and sometimes outright hostility, that historically greeted neural network researchers. This antagonism originated from a fundamental difference between two schools of thought. Traditional 'symbolic AI' researchers assumed that 'general intelligence' could be programmed into a computer using a set of logical rules. When Sejnowski asked these researchers why a traditional AI system could not see, he was told it was 'because we haven't written the vision program yet' (anyone with even minimal experience of programming knows this is laughable). In contrast, neural network researchers tried to build detailed computer models of the only intelligent machine in existence: the brain. In such a clash of ideologies, there could only be one winner. However, it took some decades for the dominant traditional AI approach to run its course, and for artificial neural networks to finally prove their worth. At least some of the delay in realizing the potential of early research in neural networks was due to a lack of raw computing power, which Moore's law inevitably delivered. Once this power became commonplace, the internet provided huge amounts of data to train ever-larger neural networks.

Sejnowski does an excellent job of describing the evolution of artificial neural networks, supported by a glossary, notes and references. However, readers looking for technical details will have to be satisfied with glimpses embodied in explainer boxes scattered throughout the text. Books that provide such details are in no short supply, and this book was never intended to be a mathematical treatise on neural networks. Instead, this is a book about how neural networks were developed by a relatively small group of (often widely scattered) individuals over several decades. It is a book about how they persisted because they believed that detailed models of brain-like neural networks could learn to solve problems that defeat the best hand-programmed computer systems. Their persistence paid off, to the extent that deep learning neural networks, when combined with reinforcement learning algorithms, have achieved superhuman levels of performance on certain tasks. Sejnowski's obvious excitement at witnessing the progress of AI makes for an engaging read. In particular, his account of how a deep learning network called AlphaGo

defeated the best human players at the game of Go, and of how another neural network (AlphaGo Zero) then taught itself to beat AlphaGo, is both astounding and fascinating.

For each new development in artificial neural networks, Sejnowski explains how it is analogous to the structure or functioning of the brain. Such analogies, amply justified in some cases, are inevitably more speculative than accounts of the neural networks themselves. However, it is no coincidence that step changes in neural network progress were usually accompanied by innovations inspired by particular design features of the brain. For example, Sejnowski relates how the invention of convolutional neural networks by Yann LeCun was inspired by the image processing known to occur in the retina.

Throughout, the writing style is fluid and insightful, and Sejnowski's deep passion for his scientific quest is never in doubt. In the final chapter, Sejnowski considers the big questions regarding the logic of life, evolution and intelligence — questions that culminate with the the exquisitely Darwinian, "What are the cost functions in nature?"

The discussions on the future impact of deep learning networks on society are thought-provoking, but less compelling than Sejnowski's accounts of its past history and present achievements. Nonetheless, the future impact of AI is a major topic of debate in the wider community, where inexpert hyperbole abounds, so the thoughtful analysis provided by a world expert like Sejnowski is urgently needed. *The Deep Learning Revolution* is an important and timely book, written by a gifted scientist at the cutting edge of the AI revolution. Historically, AI has had more than its fair share of false prophets, who confidently declared that it would become a reality within ten years. In this respect, AI can be compared to generating electricity from nuclear fusion: inevitable, but with a delivery date fixed permanently at about 20 years in the future. If Sejnowski is right, the delivery date of AI is much, much sooner than that. □

Reviewed by James V. Stone

University of Sheffield, Sheffield, UK. His book *Deep Neural Networks* will be published in March 2019.
e-mail: j.v.stone@sheffield.ac.uk

Published online: 7 January 2019
<https://doi.org/10.1038/s42256-018-0012-1>