JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 2, October 2005

Book Review

On Intelligence

Jeff Hawkins with Sandra Blakeslee, Times Books, New York, 2004, ISBN: 0-8050-7456-2, 261 Pages, $25

"On Intelligence" by Jeff Hawkins with Sandra Blakeslee, is an attempt to understand the workings of the brain and human intelligence. It presents a new perspective on how the brain (and specifically the neocortex) creates intelligence and what approach one could follow when building an intelligent system. Jeff Hawkins is an entrepreneur and computer expert, and is known for inventing handheld computers and cell phones, such as the Palm Pilot and the Treo. He is also interested in understanding the workings of the human brain and building intelligent machines. He is not an expert in artificial intelligence, however, and the reader should keep this in mind while reading the book.

The book consists of eight chapters written in 233 pages, and has a companion webpage http://www.onintelligence.org/. The website hosts some source codes for simulation and also some online discussion forums. The book is intended for readers interested in understanding how human brains perform high-level cognitive tasks and may be of particular interest to researchers trying to artificially model the cognitive tasks of the neocortex on a computer. General readers will find the book interesting and stimulating and will find explanations to their many unanswered questions such as, "how does the brain create intelligence?" Throughout the book several examples and thought experiments are presented which help assimilating the proposed ideas of the author. Many of the ideas presented in the book are hypotheses, and will need to be tested and discussed further.

The book begins with some background on why previous attempts at understanding intelligence and building intelligent machines have failed, but the meaning of intelligence itself is somewhat controversial. The author is critical of traditional Artificial Intelligence and Neural Networks as they both do not follow the "model" of our brain. The author considers the brain to be a large memory system which makes predictions based on previously stored memories from past events. According to the author, AI can be used to do some very specific tasks but it still cannot be used to build "truly intelligent" machines. Intelligence as defined by the author is "the capacity of the brain to predict the future by analogy to the past." The reviewer thinks that it would be unfair to the AI community to completely neglect the contributions made by it, given the fact that the definition of Intelligence itself is debatable. The author argues that Neural Networks did seem to offer a generalist approach but they did not satisfy three criteria that he places for a system to be truly intelligent which are the temporal aspects, the importance of feedback (even in forward propagation), and the hierarchical memory architecture of the brain.

In Chapter 3, the author discusses in brief the structure of the brain and the neocortex—the wrinkled outer layer of the brain—which he believes is primarily responsible for intelligence. He considers the neocortex to be the primary area where all the essential aspects of intelligence (e.g. perception, language, imagination, mathematics, art, music, and planning) occur. The neocortex is the most recently evolved part of the brain, and is often believed to be responsible for evolution of intelligence. The author states, "Humans are smarter, because our neocortex, relative to body size, covers a larger area, not because our layers are thicker or contain some special class of smart cells". He compares the human neocortical sheet to the neocortical sheets of a monkey and a rat. It was surprising that the author provides no comparison to the brain of a Dolphin which has a more convoluted brain than even humans.[1] A tabular comparison between the important properties of the human brain and other relatively less intelligent mammals would have been useful.

It has long been known by neuroscientists that certain regions of the cortex perform some particular mental functions. Felleman and Essen[2] have studied a detailed map of the monkey cortex. According to them, we can consider the monkey cortex to consist of many regions connected together in a complex hierarchy. Note that the

Received 9 June 2005; revision received 26 July 2005; accepted for publication 28 July 2005.

hierarchy referred to here is functional not physical. The author assumes that similarly we can consider the human neocortex to consist of certain regions arranged in a complex functional hierarchy. Mountcastle[3] pointed out that the neocortex is remarkably uniform in appearance and structure, and concluded that there is a common function, a common algorithm that is performed in all the cortical regions. The author supports Mountcastle's view and presents several examples and scientific facts to bolster his claim. It is further claimed that all the information entering the brain comes in as spatial and temporal patterns. This is one of the most important concepts presented in this chapter which later forms a basis to support author's idea of the "memory-prediction framework".

Chapter 4 discusses how memory is stored in our brains, and how they are recalled so quickly. According to the author this is made possible by "auto-associative" memory and what he calls "invariant representations". Some work has previously been done in the field of artificial auto-associative memory particularly in the field of neural networks, but with limited success. An auto-associative memory is one which can recall complete patterns given only partial or distorted inputs. It is well known that artificial auto-associative memories fail to recognize patterns if they are moved, rotated, rescaled or transformed. Our brains, however can handle these variations with relative ease. The author claims they are able to do so as they store patterns in an "invariant representation". Numerous experiments and examples are presented in order to backup this claim, although solid proof is lacking.

The next chapter introduces the "memory-prediction framework", which is the core idea of the author's theory. The author builds upon his theory by using concepts presented in earlier chapters. He discusses the work of Turing[4] and Searle[5] briefly. According to the author, prediction, not behavior, is the proof of Intelligence. He claims that intelligence is measured by the capacity to remember and predict patterns in the world, including language, mathematics, physical properties of objects, and social situations. The sequence of events involved in the learning process are that the brain receives patterns from outside world, stores them as memories, and makes predictions by combining what it has seen before and what is happening now.

In Chapter 6, which is the longest chapter in the book, the author attempts to explain in detail how the brain physically implements the memory-prediction framework. The reviewer found this chapter to be the most demanding and the ideas presented to be very speculative and debatable. It will be a challenge to readers with no previous background in neuroscience. Later in the chapter, the author rightfully admits that many of the ideas presented on how the neocortex works may prove to be wrong. He hopes that the core ideas would be preserved though the details may change in the face of new data and understanding.

In the next chapter socio-philosophical implications of the theory are considered. Many debating questions such as, "Can we train ourselves to be more creative?", and "What is consciousness?" are considered. It is an interesting question whether animals are intelligent. The author proposes two major reasons why humans are more intelligent than other mammals. Firstly, the human neocortex is large (e.g. compared to a rat or a monkey). A large neocortex allows us to learn more deep analogies, using a more hierarchical structure, and thus the ability to think more deeply than any other mammal. Secondly, humans were able to develop language. The development of language required a large neocortex capable of handling the complex system of semantics and syntax. The Reviewer failed to understand that why neural interconnections (particularly feedback) were not considered as an important reason for greater human intelligence. The book ends with a discussion on the future of intelligence, the need to build intelligent machines, and the possibility of whether we will be able to build these machines in the near future. The author, in a series of musings, presents his thoughts on what the future intelligent machines will be like, and how one can attempt to build them.

Finally, the author, in a thoughtful epilogue hopes to attract young engineers and researchers to study the neocortex, adopt the memory-prediction framework, and build intelligent machines. He hopes that if the human neocortical algorithm could be successfully applied in an artificial device with artificial senses, it would be possible to construct intelligent machines. Though, it might not be possible to replicate an entire human being, intelligent machines with specialized tasks seem possible in near future.

Overall, the book is very well written, organized, and interesting to read. All the chapters (except may be chapter 6) have been written for a general audience, and do not attempt to summarize previous work in AI. A series of "Testable predictions" are presented as an Appendix to verify the validity of the author's new ideas. It is admitted by the author that even if these predictions are shown to be true, they are not sufficient to prove the author's hypotheses, but it would be strong evidence in support of them.

This book is not intended for readers looking for a philosophical approach towards understanding the brain; they should consider Searle[5] instead. This book is also not for readers expecting a complete scientific theory with proof on how the brain works; instead it presents the author's logical views (from the viewpoint of a computer scientist) along with examples and scientific proofs wherever possible which might help in understanding the working of our brain. The reader might want to also read Newell.[6] Some additional texts on Artificial Intelligence and Machine Learning may also provide supplemental reading. Readers looking for a book in Artificial Intelligence should consider Russell and Norvig,[7] Winston,[8] or Jackson.[9] Readers interested in machine learning algorithms should consider Mitchell[10] instead. There are other interesting, and may be controversial, books by Kurzweil,[11] Moravec,[12] and Dennett.[13] There has also been a great deal of work on cognitive architectures that should be examined to fully understand the other work in this field, for example EPIC,[14] SOAR,[15] and ACT-R.[16]

While this book is not a rigorous scholarly work, it might be valuable as a vehicle to get people (especially young people) excited about intelligent systems. It might also help to stimulate additional discussions about artificial intelligence.

Acknowledgment

I would like to sincerely thank Prof. Lyle N. Long of The Pennsylvania State University for his support, encouragement and many useful discussions.

References

[1]Howard, Carol J., "Dolphin Chronicles," Bantam Books, NY, 1995.

[2]Felleman, D. J., and Van Essen, D. C., "Distributed Hierarchical Processing in the Primate Cerebral Cortex," Cerebral Cortex, vol. 1 (Jan/Feb 1991), pp. 1–47.

[3]Mountcastle, V. B., "An organizing principle for cerebral function: The unit module and the distributed system," in Edelman, G.M., Mountcastle, V.B., The Mindful Brain, Cambridge: MIT, 1978, pp. 1–50.

[4]Turing, A. M., "Computing Machinery and Intelligence," Mind, vol. 59(1950), pp. 433–60.

[5]Searle, John R., "Mind: A Brief Introduction," Oxford University Press, New York, 2004.

[6]Newell, A., "Unified Theories of Cognition," Harvard University Press, 1994.

[7]Russell, S., and Norvig, P., "Artificial Intelligence: A Modern Approach," Pearson Education Inc., New Jersey, 2003.

[8]Winston, P. H., "Artificial Intelligence," Addison Wesley, 1992.

[9]Jackson, P., "Expert Systems," Addison Wesley, 1999.

[10]Mitchell, T. M., "Machine Learning," McGraw-Hill, 1997.

[11]Kurzweil, R., "The Age of Spiritual Machines," Penguin, NY, 1999.

[12]Moravec, H., "Robot: Mere Machine to Transcendent Mind," Oxford University Press, 2000.

[13]Dennett, D. C., "Consciousness Explained," Back Bay Books, 1992.

[14]Laird, J. E., Newell, A., and Rosenbloom, P. S. "Soar: An architecture for general intelligence," Artificial Intelligence, 33(1): 1–64, 1987.

[15]Meyer, D. E., & Kieras, D. E., "A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms," Psychological Review, 104(1), 3–65, 1997.

[16]Anderson, J. R., & Lebiere, C., "The atomic components of thought," Mahwah, NJ, Erlbaum, 1998.

Ankur Gupta
Graduate Research Assistant
Department of Aerospace Engineering
The Pennsylvania State University
University Park, PA 16802, USA