

On Intelligence

From Wikipedia, the free encyclopedia

On Intelligence: How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines is a book by Palm Pilot-inventor Jeff Hawkins with *New York Times* science writer Sandra Blakeslee. The book explains Hawkins' memory-prediction framework theory of the brain and describes some of its consequences. (Times Books: 2004, ISBN 0-8050-7456-2)

Contents

- 1 Outline
- 2 A personal history
- 3 The theory
- 4 Predictions of the theory of the memory-prediction framework
 - 4.1 Enhanced neural activity in anticipation of a sensory event
 - 4.2 Spatially specific prediction
 - 4.3 Prediction should stop propagating in the cortical column at layers 2 and 3
 - 4.4 "Name cells" at layers 2 and 3 should preferentially connect to layer 6 cells of cortex
 - 4.5 "Name cells" should remain ON during a learned sequence
 - 4.6 "Exception cells" should remain OFF during a learned sequence
 - 4.7 "Exception cells" should propagate unanticipated events
 - 4.8 "Aha! cells" should trigger predictive activity
 - 4.9 Pyramidal cells should detect coincidences of synaptic activity on thin dendrites
 - 4.10 Learned representations move down the cortical hierarchy, with training
 - 4.11 "Name cells" exist in all regions of cortex
- 5 See also
- 6 References
- 7 External links
- 8 Reviews

On Intelligence



Front Cover

Author	Jeff Hawkins & Sandra Blakeslee
Country	 United States
Language	English
Subject	Psychology
Publisher	Times Books
Publication date	2004
Media type	Paperback
Pages	272
ISBN	0-8050-7456-2
OCLC	55510125 (http://worldcat.org/oclc/55510125)
Dewey Decimal	612.8/2 22
LC Class	QP376 .H294 2004

Outline

Hawkins outlines the book as follows:

The book starts with some background on why previous attempts at understanding intelligence and building intelligent machines have failed. I then introduce and develop the core idea of the theory, what I call the memory-prediction framework. In chapter 6 I detail how the physical brain implements the memory-prediction model—in other words, how the brain actually works. I then discuss social and other implications of the theory, which for many readers might be the most thought-provoking section. The book ends with a discussion of intelligent machines—how we can build them and what the future will be like. (p. 5)

A personal history

The first chapter is a brief history of Hawkins' interest in neuroscience juxtaposed against a history of artificial intelligence research. Hawkins uses a story of his failed application to the Massachusetts Institute of Technology to illustrate a conflict of ideas. Hawkins believed (and ostensibly continues to believe) creating true artificial intelligence will only be possible with intellectual progress in the discipline of neuroscience. Hawkins writes that the scientific establishment (as symbolized by MIT) has historically rejected the relevance of neuroscience to artificial intelligence. Indeed, some artificial intelligence researchers have "[taken] pride in ignoring neurobiology" (p. 12).

Hawkins is an electrical engineer by training, and a neuroscientist by inclination. He used electrical engineering concepts as well as the studies of neuroscience to formulate his framework. In particular, Hawkins treats the propagation of nerve impulses in our nervous system as an encoding problem, specifically, a future predicting state machine, similar in principle to feed-forward error-correcting state machines.

The theory

Main article: Memory-prediction framework

Hawkins' basic idea is that the brain is a mechanism to predict the future, specifically, hierarchical regions of the brain predict their future input sequences. Perhaps not always far in the future, but far enough to be of real use to an organism. As such, the brain is a feed forward hierarchical state machine with special properties that enable it to learn.

The state machine actually controls the behavior of the organism. Since it is a feed forward state machine, the machine responds to future events predicted from past data.

The hierarchy is capable of memorizing frequently observed sequences (Cognitive modules) of patterns and developing invariant representations. Higher levels of the cortical hierarchy predict the future on a longer time scale, or over a wider range of sensory input. Lower levels interpret or control limited domains of experience, or sensory or effector systems. Connections from the higher level states predispose some selected transitions in the lower-level state machines.

Hebbian learning is part of the framework, in which the event of learning physically alters neurons and connections, as learning takes place.

Vernon Mountcastle's formulation of a cortical column is a basic element in the framework. Hawkins places particular emphasis on the role of the interconnections from peer columns, and the activation of columns as a whole. He strongly implies that a column is the cortex's physical representation of a state in a state machine.

As an engineer, any specific failure to find a natural occurrence of some process in his framework does not signal a fault in the memory-prediction framework *per se*, but merely signals that the natural process has performed Hawkins' functional decomposition in a different, unexpected way, as Hawkins' motivation is to

create intelligent machines. For example, for the purposes of his framework, the nerve impulses can be taken to form a temporal sequence (but phase encoding could be a possible implementation of such a sequence; these details are immaterial for the framework).

Predictions of the theory of the memory-prediction framework

His predictions use the visual system as a prototype for some example predictions, such as Predictions 2, 8, 10, and 11. Other predictions cite the auditory system (Predictions 1, 3, 4, and 7).

- An Appendix of 11 Testable Predictions:

Enhanced neural activity in anticipation of a sensory event

1. In all areas of cortex, Hawkins (2004) predicts "we should find *anticipatory cells*", cells that fire in anticipation of a sensory event.

Note: As of 2005 mirror neurons have been observed to fire *before* an anticipated event.^[1]

Spatially specific prediction

2. In primary sensory cortex, Hawkins predicts, for example, "we should find anticipatory cells in or near V1, at a precise location in the visual field (the scene)". It has been experimentally determined, for example, after mapping the angular position of some objects in the visual field, there will be a one-to-one correspondence of cells in the scene to the angular positions of those objects. Hawkins predicts that when the features of a visual scene are known in a memory, anticipatory cells should fire *before* the actual objects are seen in the scene.

Prediction should stop propagating in the cortical column at layers 2 and 3

3. In layers 2 and 3, predictive activity (neural firing) should stop propagating at specific cells, corresponding to a specific prediction. Hawkins does not rule out anticipatory cells in layers 4 and 5.

"Name cells" at layers 2 and 3 should preferentially connect to layer 6 cells of cortex

4. Learned sequences of firings comprise a representation of *temporally constant invariants*. Hawkins calls the cells which fire in this sequence "name cells". Hawkins suggests that these *name cells* are in layer 2, physically adjacent to layer 1. Hawkins does not rule out the existence of layer 3 cells with dendrites in layer 1, which might perform as *name cells*.

"Name cells" should remain ON during a learned sequence

5. By definition, a *temporally constant invariant* will be active during a learned sequence. Hawkins posits that these cells will remain active for the duration of the learned sequence, even if the remainder of the cortical column is shifting state. Since we do not know the encoding of the sequence, we do not yet know the definition of *ON* or *active*; Hawkins suggests that the ON pattern may be as simple as a simultaneous AND (i.e., the name cells simultaneously "light up") across an array of name cells.

See Neural ensemble#Encoding for *grandmother neurons* which perform this type of function.

"Exception cells" should remain OFF during a learned sequence

6. Hawkins' novel prediction is that certain cells are inhibited during a learned sequence. A class of cells in layers 2 and 3 should NOT fire during a learned sequence, the axons of these "exception cells" should fire

only if a local prediction is failing. This prevents flooding the brain with the usual sensations, leaving only exceptions for post-processing.

"Exception cells" should propagate unanticipated events

7. If an unusual event occurs (the learned sequence fails), the "exception cells" should fire, propagating up the cortical hierarchy to the hippocampus, the repository of new memories.

"Aha! cells" should trigger predictive activity

8. Hawkins predicts a cascade of predictions, when recognition occurs, propagating down the cortical column (with each saccade of the eye over a learned scene, for example).

Pyramidal cells should detect coincidences of synaptic activity on thin dendrites

9. Pyramidal cells should be capable of detecting coincident events on thin dendrites, even for a neuron with thousands of synapses. Hawkins posits a temporal window (presuming time-encoded firing) which is necessary for his theory to remain viable.

Learned representations move down the cortical hierarchy, with training

10. Hawkins posits, for example, that if the inferotemporal (IT) layer has learned a sequence, that eventually cells in V4 will also learn the sequence.

"Name cells" exist in all regions of cortex

11. Hawkins predicts that "name cells" will be found in all regions of the cortex.

See also

- Hierarchical temporal memory, a technology by Hawkins's startup Numenta Inc. to replicate the properties of the neocortex.
- Memory-prediction framework

References

1. ^ Fogassi, Leonardo, Pier Francesco Ferrari, Benno Gesierich, Stefano Rozzi, Fabian Chersi, Giacomo Rizzolatti (April 29, 2005). "Parietal lobe: from action organization to intention understanding" (<http://www.unipr.it/arpa/mirror/pubs/pdffiles/Fogassi-Ferrari2005.pdf>) (PDF). *Science* **308** (5722): 662–667. doi:10.1126/science.1106138 (<http://dx.doi.org/10.1126%2Fscience.1106138>). PMID 15860620 (<http://www.ncbi.nlm.nih.gov/pubmed/15860620>). Retrieved 2006-11-18.

External links

- OnIntelligence.com (<http://www.onintelligence.com>) - official website
- A Hierarchical Bayesian Model of Invariant Pattern Recognition in the Visual Cortex (<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.6744&rep=rep1&type=pdf>), a conference paper by George & Hawkins
- Saulius Garalevicius' research page (<http://www.phillylac.org/prediction/>) - Research papers and programs presenting experimental results with Bayesian models of the Memory-Prediction Framework
- Project Neocortex (<http://sourceforge.net/projects/neocortex/>) - An open source project for modeling

Reviews

- Machine Intelligence Meets Neuroscience (<http://www.computer.org/computer/homepage/0105/random/index.htm>) (By Bob Colwell, published in IEEE's *Computer*, January 2005)
 - Above link is broken, although the Internet Archive has a copy (<http://web.archive.org/web/20050204164903/http://www.computer.org/computer/homepage/0105/random/index.htm>).
 - Machine Intelligence Meets Neuroscience (<http://ieeexplore.ieee.org/iel5/2/30112/01381247.pdf>) (full text for IEEE Explore's online subscribers)
 - Machine Intelligence Meets Neuroscience (<http://csdl2.computer.org/persagen/DLAbstoc.jsp?resourcePath=/dl/mags/co/&toc=comp/mags/co/2005/01/r1toc.xml&DOI=10.1109/MC.2005.24>) (citation only)
- A review by Franz Dill (http://future.iftf.org/2004/10/jeff_hawkins_on.html)
- On Intelligence, People and Computers (<http://www.techcentralstation.com/article.aspx?id=112204B>) (Arnold Kling, Tech Central Station, 22 November 2004)
- On Biological and Digital Intelligence (<http://www.goertzel.org/dynapsyc/2004/OnBiologicalAndDigitalIntelligence.htm>) A review by Ben Goertzel (7 Oct 2004)

Retrieved from "http://en.wikipedia.org/w/index.php?title=On_Intelligence&oldid=592633952"

Categories: [2004 books](#) | [Science books](#) | [Artificial intelligence publications](#)

| [Books about human intelligence](#)

- This page was last modified on 27 January 2014 at 13:30.
- Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy.
Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.