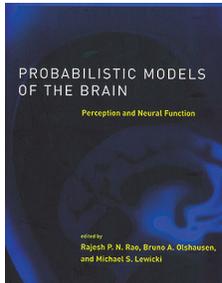# Brain models: the next generation

Probabilistic Models of the Brain

edited by Rajesh P.N. Rao, Bruno Olshausen and Michael Lewicki
MIT Press, Cambridge, Massachusetts, 2002. $50.00
hardcover, pp 334
ISBN 0-262-18224-6

Reviewed by Josh McDermott

The notion that the brain performs inferences dates back at least to Helmholtz, who observed that most perceptual problems are inherently ill posed—our sensory input vastly underdetermines the structure of the world that we perceive. The brain must therefore resort to making a guess as to what lies out in the world, based on prior knowledge of what the world is like. Helmholtz coined the term unconscious inference to refer to the myriad perceptual processes that faithfully operate beneath our awareness, and his ideas form the foundation of modern perceptual science.

Despite widespread acceptance of the idea that perception involves inference, probabilistic models of these inferences have only recently entered mainstream neuroscience and psychology. Early neural network models, for instance, were not described in the language of probability, although many of them, in retrospect, have probabilistic interpretations. But given the uncertainty present in everything the brain does, both due to the ill-posed nature of its tasks and to internal and external sources of noise, probabilities are unavoidable. Thus it is now standard for models of the brain to make probabilities explicit.

Probabilistic Models of The Brain, edited by Rao, Olshausen and Lewicki, surveys this next generation of models. The book, which grew out of a 1998 NIPS workshop on the same topic, consists of 16 invited chapters written by leaders in computational vision and neuroscience. It is split into two parts. The first deals with relatively abstract models of what the brain does, and the second

Josh McDermott is in the Department of Brain and Cognitive Sciences, MIT, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.
e-mail: jhm@mit.edu

covers brain models that explicitly involve neurons. This division reflects a distinction made over twenty years ago by David Marr, who is often considered the founder of computational vision. Marr famously distinguished between characterizing the neural circuitry carrying out a particular computation, and characterizing the computation itself (in other words, the rule that describes the output of the system as a function of its input, independent of the neural circuitry that implements it). Marr contended that the computations being implemented by the brain are often far from obvious, and that a correct theory of the computation would place important constraints on what neural circuits might be implementing it. Despite Marr's influence, models at this level are often neglected in computational neuroscience circles; indeed, the term computational has for the most part lost any reference it once had to Marr, and generally just refers to anything involving math. It is therefore a pleasant surprise to see that the first half of the book is devoted to models at this more abstract level of analysis.

The book begins by taking examples from vision to show how some of the brain's perceptual computations may embody principles of Bayesian statistics, in which prior knowledge about the world can be optimally combined with sensory input. Other chapters use information theory to determine limitations of bottom-up theories of perception. Also included are chapters describing the use of Markov networks and support vector machines, two current focuses of the machine learning community, to model different aspects of vision. All these approaches shed light on what the brain is computing, without considering how these computations might be implemented in networks of neurons.

The second half of the book complements the first by focusing on models in

which neurons are an explicit component, often showing how neural networks may instantiate particular computational principles. For instance, Simoncelli and colleagues show that adaptation effects in cortical neurons are predicted by contrast normalization mechanisms if the normalization parameters are set to make the resulting neural responses as independent as possible. This provides new support for Horace Barlow's well-known proposal that one goal of cortical computation is to produce neural responses that are statistically independent. Similarly, Olshausen shows that if neural responses to natural movies are forced to be sparse and independent, the receptive fields that result resemble those found physiologically in V1. These two lines of research provide paradigmatic examples of models providing support for theories where the relevant experimental evidence is hard to come by; experimental tests of independence would require multielectrode recordings, which are relatively uncommon due to their technical difficulty.

Other chapters explore general mechanisms of neural computation. Topics discussed include the representation of probability distributions by spiking neurons, the computational utility of cortical feedback, and a new look at the possible role of synchronous spikes.

Overall, the book provides an eclectic sample of state-of-the-art computational neuroscience, and is laudable for its scope. Evident in this book is a problem commonly encountered in computational neuroscience: the tools of the trade are not routinely taught in traditional neuroscience and psychology programs, which often makes it hard for outsiders to understand the ideas at the heart of the research. This will be a problem in this book as well, but the standard of writing is high, and several of the papers are particularly accessible. If there is a weakness of the book, it is that there is relatively little effort to convince non-computational neuroscientists of the need for probabilistic models. A few of the chapters do seem specifically tailored to the non-specialist, perhaps most notably, the introduction to Bayesian inference by Mamassian, Landy and Maloney. But even this chapter, though providing a lucid explanation of Bayesian modeling, may leave the skeptic unsure as to why one ought to formulate such models in the first place. However, perhaps the best path for the skeptic is to read about some of the best current examples of modeling successes, and this book certainly provides that opportunity.