REVIEW

The Model Thinker by Scott E. Page

TED G. LEWIS
Email: tedglewis@icloud.com

The Model Thinker by Scott E. Page is a courageous attempt to boil-down a handful of major mathematical models into their essence. It distills both traditional models such as linear regression and game theory as well as "modern" complexity theory models developed over the past 40 years based largely on network science and non-linear systems theory. Page's motivation is to fill the gap between his YouTube.com lectures, which he practices in front of his dog, and the written word. I say it is a courageous attempt because the models cover a broad swath of disciplines and topics, which he tries to make accessible to the semi-mathematical reader. Like a good sauce, sometimes the topics are simmered too long, and sometimes they need more simmering. Nonetheless, he has produced an encyclopedia of important models. This is not easy to do and his attempt is commendable.

According to Page, we all need to become many-model thinkers. The first 3 chapters define a many-model thinker as someone who considers many models when modeling reality. Dogmatically sticking with one model often leads to the wrong answer. Instead, Page advocates, "the application of ensembles of models to understand complex phenomena." There should be no argument with this goal, but I think it takes up entirely too much of the book.

I became uncomfortable with these preliminary chapters when Information was defined as "names and partitions [of] data into categories." I was taught that Information equals entropy. This misdirection is rectified in chapter 12, where he discusses Shannon's theory, but I found it distracting. Page claims data is transformed into information, and information is transformed into knowledge, which leads to wisdom. These form a *wisdom hierarchy*, and the job of the modeler is to traverse this hierarchy. Personally, I found these chapters unnecessary. They delayed my access to the rich set of models that make up most of the book.

The book needs editing. Grammar and meaning are totally lost in sentences like, "The core idea is that many-model thinking produces wisdom through a crowd of a diverse logical frames (p. 1)." Or, how about, "Given that the brain differs at each level, we need multiple models, and those models differ (p. 11)." It doesn't improve later on, when grammatical chaos strikes again on p. 32, "Attempts to constructing a collection of diverse, accurate models can encounter similar problems."

After a rough start, the book gets down to business in chapter 5: *Normal Distributions*. Hundreds of books and articles have been written about how the Normal Distribution is the wrong distribution to use to describe the highly connected and complex world. Page uses it like the rest of us, to explain why long-tailed distributions are better. He sets us up for chapter 6: Power Law Distributions. This is where the influence of the Santa Fe Institute begins to show. Unlike many authors writing about the long tail, Page successfully identifies the root cause of long tails— the lack of independence among measurements made along the x-axis. In a complex hyper-connected world, almost everything is dependent on something else. In other words, chance is conditional. This dependence leads to complex behaviors like preferential attachment and other forms of self-organization. I was a bit disappointed that Page does not give credit where credit is due, for example, Per Bak created the self-organized criticality model that underlies so much of complexity theory. One can even go back further and recognize economist Friedrick Hayek's concept of spontaneous order, which influenced Bak and others. Old ideas often re-appear in new forms, and a book like this should show how great ideas emerge and flow together.

The next three chapters cover standard linear models, non-linear S-shaped adoption models, and game theory models in workman like fashion. Chapters 18 (System Dynamics), 19 (Feedback), and 22 (Cooperation) complete the topics needed for students from across many disciplines. Chapter 10: *Network Models*, introduces the reader to the new discipline of complex networks. The coverage is some-

what dated, which impacts the treatment of contagions in networks in the following chapter. An updated chapter on networks should include spectral analysis and note the use of eigenvalues as a measure of centrality in search algorithms such as Google's original PageRank search algorithm.

Chapter 11: Broadcast, Diffusion, and Contagion should also describe the use of spectral radius in place of degree centrality to explain how contagions spread in networks without ideal scale-free, small world, or random topologies. Real world networks are not pure. Rather, they are combinations of hubs, clusters, and random connections. Spectral radius is more general and applies to mixed-topology networks. Similarly, a modern treatment of "nearest-neighbor contagions" would be appreciated, because they more accurately describe the spread of misinformation known as "fake news"— a topic of interest over the past few years.

Page takes us on a journey through a wide variety of topics, which is one of the book's most impressive features. This voyage begins in earnest with chapter 20 (Spatial and Hedonic Choice), which is important to the study of politics and economics. The lack of careful editing shows up again, however, on p. 229, where the word "six" should apparently be "eight" to match Figure 20.1. Typos persist on the next page where the Spatial Competition Model defines "C>", instead of "C>0."

Chapters 13, 17, and 26 threaten to establish the foundations of deep learning in artificial intelligence (AI). Random walks and Markov chains (Markov clustering) are common tools used in classifiers, e.g., convolutional neural networks. I think the emphasis of pure mathematics versus algorithms exposes one of the deficiencies of the book. By restricting models to mathematics versus computer algorithms and heuristics, Page overlooks a segment of tools useful for many-model thinking. Not every model is amenable to pure math. The book would be much more useful and powerful if it incorporated algorithms as well as mathematics. Connecting Markov chains, learning, and classification with AI would make the book more appealing to the modern reader.

Chapters 23 through 25 are important in decision analysis and could be expanded into an entire book of its own. Remaining chapters cover a broad spectrum of fields and provide numerous examples from politics, economics, and drug addition. It is a most ambitious undertaking.

Finally, there is no index. It is obvious from the extensive Notes and Bibliography sections that the author has thoroughly researched his topics, hence more the pity that the publisher did not include an index. An index and good editor (to fix grammar and clarity) would greatly enhance this important book. I put my copy on the shelf marked "reference," and expect to use it more often than Google for finding models.