

A View of Inflation—Eternally Going on Somewhere in the Multiverse

Our Cosmic Habitat

► Martin Rees
Princeton U. Press, Princeton, N.J., 2001. \$22.50 (205 pp.). ISBN 0-691-08926-4

Reviewed by Michael S. Turner

Astronomy is in the midst of an explosive period of discovery and deepening understanding, made possible by new technology and new ideas. Just 100 years ago the universe was dated by astronomers to be 30 million years old and consisted of our galaxy, with only a million of its 100 billion stars observed. The Solar System was known to have only 7 planets, and the workings of stellar furnaces were a mystery. The largest telescope was a 40-inch refractor that used photographic plates to capture 1% of the incident cosmic light.

Today, the largest optical telescope is 400 inches in diameter and charge-coupled devices capture nearly 100% of the incident light. Our eyes on the universe span from radio to TeV gamma rays and also include cosmic rays with energies up to 10^{21} eV, MeV neutrinos from the Sun and a supernova, and, possibly soon, gravity waves. The Sloan Digital Sky Survey has cataloged close to 100 million galaxies, including four quasars whose light was emitted when the universe was less than one-seventh its present size and less than a billion years old. Four independent methods peg the age of the universe at 14 ± 1 billion years. Known extrasolar planets outnumber the planets in our own solar system by 10 to 1.

In short, science can now tell the story of the universe from quark soup to the emergence of life on the third rock from the Sun. And no one is more capable of doing so than Sir Martin Rees, leading British astrophysicist and Astronomer Royal. He is articulate, witty, and authoritative, and he has the intellectual breadth to cover this enormous range in space, time, and subject matter.

MICHAEL S. TURNER is the Rauner Distinguished Service Professor at the University of Chicago and a staff scientist at Fermi National Accelerator Laboratory.

In 2000, Rees delivered the first series of Scribner Lectures at Princeton University; they are the substance of his most recent book, *Our Cosmic Habitat*. The book's style is less formal than the usual popular science book, but I think it works very well, for the generalist and even for the specialist.

The title of the prologue is the question raised by Einstein, "Could God have made the world any differently?" It sets the stage. This question, which in Einstein's time was certainly outside the realm of science, may now be within its purview. What makes this so, according to Rees, is the concept of the multiverse: a universe consisting of infinitely many causally distinct regions.

To prepare for the multiverse, Rees takes the reader on a wonderful tour of our universe, from its Big-Bang beginning to stars, planets, and life; he always clearly differentiates what we know with reasonable confidence from current fad and wild speculation.

In the final part of the book, Rees turns to the multiverse, which has its origins in inflationary universe theory. According to inflation, all that we see today traces back to a small bit of the pre-inflationary universe that underwent a burst of exponential expansion. Alex Vilenkin, Andrei Linde, and others have shown that inflation is eternal—if it ever occurred, it is occurring now, it occurred in the past, and it will occur in the future. Different inflationary bubbles are causally distinct, which leads to a multiverse structure for the universe. If true, that is a realization as profound as Copernicus's banishing of Earth from the center of the universe.

But here many cosmologists (including me) get conflicted. As profound as the multiverse concept may be, if it isn't testable, it isn't science. Although we are well on our way to testing the idea that our portion of the universe originated from a burst of inflationary expansion, the multiverse seems beyond the reach of similar direct testing, because, according to our current understanding, the different bubbles cannot communicate. However, Rees argues that the multiverse may even be testable now—I leave readers of the

book to decide whether or not his arguments are persuasive.

What Rees finds so attractive about the multiverse is that it gets at Einstein's question, or at least Rees's reformulation of it. In a previous successful book, *Just Six Numbers* (Basic Books, 2000), Rees identified a set of six numbers that had to be "just so" in order to result in a universe like ours. (The six numbers include the level of primeval lumpiness in the dark matter, the weakness of gravity, the binding energy of atoms, the amount of ordinary matter, and the strength of the weak forces.) Rees's version of Einstein's question is then, Could these six numbers have been chosen differently?

Marrying the multiverse to modern ideas in particle physics leads to the possibility of finding in the different bubbles, different local "bylaws" such as the local values for Rees's six numbers, the suite of stable particles, and even numbers of large dimensions. The existence of our universe comes not from any special choice but from the large number of tries. Fortunately, one can only imagine Einstein's reaction, as Rees has taken dice rolling to a new cosmic level.

Our Cosmic Habitat is science writing at its best: An eminent and articulate scientist writes about the frontiers of research and is not afraid to give his own perspective on the most thought-provoking questions. Whether or not his views on the multiverse hold up is irrelevant; he has taken his readers from quark soup to life on Earth, stimulating them in the process to think about one of the truly big questions.

Bioinformatics: The Machine Learning Approach

► Pierre Baldi and Søren Brunak
MIT Press, Cambridge, Mass., 2nd ed., 2001 [1998]. \$49.95 (452 pp.). ISBN 0-262-02506-X

Bioinformatics is an amorphous discipline that could be described as "biologically inspired computer science." But bioinformatics also draws ideas

from physics, chemistry, biochemistry, mathematics, and statistics. This interesting blend of fields created a tower of Babel out of which evolved a communication currency based on the tools and techniques useful for computer analyses of biological data. It is thus not surprising that bioinformatics can be presented by means of a methodology thread, which is the way chosen by Pierre Baldi and Søren Brunak in their book *Bioinformatics: The Machine Learning Approach*.

The core of *Bioinformatics* is the Bayesian probabilistic framework. That focus permeates all the calculations and algorithms presented in the book and gives the book a strong sense of continuity and unity. It is conceivable that the book could have been organized in terms of such biologically important topics as protein structure prediction, sequence alignment, protein family classification, and the like. That the book is organized around techniques underscores the authors' intent: Theirs is a book on methods rather than on specific problems.

Baldi and Brunak cover an enormous amount of information. Here are some highlights: Chapter 1 includes an interesting discussion on the quality of data, and the sources of the many errors contained in the rapidly expanding biological databases. Chapter 2 sets the tone of the book in terms of how to "think" Bayesian. Examples of Bayesian inference are provided in chapter 3. An account of many important optimization techniques is the subject of chapter 4. Chapters 5 through 8 are the juiciest ones, introducing the book's main machine-learning algorithms: neural networks and hidden Markov models. These chapters also deal in detail with some of the key questions of bioinformatics, such as protein secondary structure prediction, intron-splice-site prediction, and identification of the important G-protein coupled receptors (GPCR) protein family. Chapter 9 confers additional unity to the book, as it presents neural networks and hidden Markov models in light of more general probabilistic graphical models. A brief, example-free chapter 10 deals with the inference of phylogenetic trees. Chapter 11 succinctly introduces stochastic grammars and linguistics application to RNA secondary structure prediction. Chapter 12 introduces a Bayesian hypothesis-testing scheme for gene-expression analysis in DNA microarray data and a brief summary of clustering techniques. Unfortunately no concrete biological example

illustrates this chapter. The last chapter (13) is a very useful list of public database resources that are accessible over the Internet. The book finishes with a set of six appendices that dig deeper into the technicalities of some of the subjects touched upon in the earlier chapters. Chapter 12 and appendix E (on support vector machines and Gaussian processes) are new and welcome additions to this second edition (in which many of the typos of the first edition have managed to survive, and new typos have crept in).

The authors aim at an audience of students and more advanced researchers with diverse backgrounds, who in the authors' view, do not need previous knowledge of DNA, RNA, and proteins. Such knowledge, however, seems to me to be a necessary prerequisite. I see this book as aimed at an audience with some bioinformatics experience and a desire to get deeper into a subset of methods used in the field. Readers with physics training may find some examples particularly easy to relate to, as many optimization results are presented in terms of associated free energies. Readers not acquainted with physics concepts, however, will probably not appreciate the many free-energy analogies.

The book is lucidly written. The basic algorithms are complemented with thoughtful comments drawn from the authors' considerable hands-on experience. Long passages of the book read like thorough reviews of the subject being presented and include extensive descriptions of the literature. Indeed, the reference list grew to 587, from the 452 references contained in the previous edition. The passages describing the authors' own work (including an analysis of the GPCR protein family and an analysis of symmetries of the genetic code) convey an excitement that goes beyond the pure presentation of a methodology.

Throughout the book, the authors argue that the use of priors (probabilities assigned to models that explain the data prior to our present experimentation) in Bayesian inference is a source of flexibility in the analysis. The lack of explicit priors was one of the criticisms the authors made (in chapter 3) of the traditional derivation of the Gibbs distribution in statistical mechanics. In the context of statistical mechanics, however, the use of statistical ensembles is indeed an indirect way to choose priors, except that these priors are not associated with a subjective belief but rather are dictated by the physics of the problem. In the micro canonical

ensemble, for example, Liouville's theorem tells us that the probability of each microstate is uniform. The use of conjugate priors is nicely illustrated in chapter 12, where the parameters of the posterior probability combine information from the prior probability and the data in a very meaningful way.

This is a very good book, written with a high level of erudition and insight. It provides an excellent account of the important place that machine learning plays in bioinformatics. It also constitutes a reference source of methodologies and applications for the computational biologist. And it should certainly increase the degree of belief in Bayesian inference, even for those readers without much prior experience.

GUSTAVO A. STOLOVITZKY

*IBM T. J. Watson Research Center
Yorktown Heights, New York*

The Lattice Boltzmann Equation for Fluid Dynamics and Beyond

▶ Sauro Succi
Oxford U. Press, New York, 2001.
\$100.00 (288 pp.).
ISBN 0-19-850398-9

The Navier–Stokes equations describe the flow of simple fluids. The equations can be derived easily by applying the laws of mass conservation and Newton's second law to an elementary fluid volume ("volumelet," in the charming terminology of Sauro Succi's *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*) incorporating the assumption that stress and strain are proportional.

Despite their apparent simplicity, the Navier–Stokes equations describe rich physics, some of which is still not understood in detail. For example, at high Reynolds numbers (low viscosities, high velocities, and large length scales) flow becomes turbulent, as exemplified by waterfalls, blood flow, and atmospheric flows. In such flows, the nonlinear nature of the Navier–Stokes equations is making itself felt; many different length scales contribute to the fluid motion.

Solving the Navier–Stokes equations preoccupies many scientists and engineers. Geophysicists need to predict the flow of ocean currents, engineers designing cars aim for streamlining to reduce turbulence, and reactor physicists model flow and heat transfer under extreme conditions. For computational fluid dynamics, commercial