

Book Review

Complexity, Entropy and the Physics of Information, edited by Wojciech H. Zurek. Proceedings Volume VIII in the Sante Fe Institute Studies in the Sciences of Complexity. New York: (Addison-Wesley, 1990, 530 pp. + xiii, \$29.25)

Reviewer: Thomas M. Cover, *Fellow, IEEE*

This excellent volume is devoted to the physics of information. For many years, John Wheeler, one of the key players in quantum theory and general relativity, has asserted that information plays a central role in physics. It was perhaps his vision, and that of Wojciech Zurek, that led to the workshop held in Santa Fe, NM, in 1989, the proceedings of which are under review here.

This book examines the relation between thermodynamics and information-theoretic entropy, algorithmic complexity, the physics of computation, information processing in chemical and biological systems, and the role of information in quantum measurement and cosmology.

A listing of some of the topics suggests the far-reaching and comprehensive ambitions of the physicists involved. The titles include, "The Entropy of Black Holes," by V. F. Mukhanov, "Some Simple Consequences of the Loss of Information in a Spacetime with a Horizon," by Shin Takagi, "Why is the Physical World so Comprehensible?" by P. C. W. Davies, "Quantum Mechanics and Algorithmic Complexity" by J. W. Barrett, "Algorithm Information Content, Church-Turing Thesis, Physical Entropy, and Maxwell's Demon" by W. Zurek, "How to Define Complexity in Physics, and Why" by Charles Bennett, "Entropy and Quantum Mechanics" by M. Hossein Partovi, and the ambitious "Quantum Mechanics in the Light of Quantum Cosmology" by Murray Gell-Mann and James Hartle.

One of the more interesting contributions to this volume is W. H. Zurek's contribution, "Algorithmic Information Content, Church-Turing Thesis, Physical Entropy, and Maxwell's Demon." Maxwell's demon sits on a trap door on a glass partition dividing a boxful of homogeneous gas. He lets the hot atoms through to the left and the cold atoms through to the right, thereby violating the second law of thermodynamics by creating a temperature differential that can then be used to do physical work. Szilard, Brillouin and Bennett, among others, argued that we had forgotten to balance the books. The state of mind of Maxwell's demon has changed because of the information he has gathered. Until we have zeroed out his memory, we have not reduced the problem to its starting state.

Zurek asks how much information has to be zeroed out. Strangely enough, instead of trying to argue that many bits of information have been gathered and that each costs $kT \ln 2$ in entropy to zero out, he allows the demon to first compress his memory using the notion of algorithmic complexity put forth by Komogorov, Solomonoff and Chaitin and then zero out that smaller quantity. Thus, in a way, he is trying to weaken Brillouin's attempt to balance the books. Zurek shows that even when the demon's information is maximally compressed, the books still balance, entropy still increases, and the second law survives. I think that it is fair to say, though, that a discussion of Maxwell's demon will continue to perplex scientists. Apparently, it has to do with whether $kT \ln 2$ is the appropriate amount of entropy associated with the erasure of one bit in a physical system.

Charles Bennett gives us an interesting discussion of how to define complexity in physics. In particular, he compares the notion of thermodynamic depth due to Seth Lloyd and Heinz Pagels, which has to do with the amount of entropy produced during a state's actual evolution, with Bennett's own notion of logical depth, which is the execution time required to generate the object in question by a nearly incompressible universal computer program. As Bennett says, "Logically deep objects, in other words, contain internal evidence of having been the result of a long computation or slow-to-simulate dynamical process and could not plausibly have originated otherwise."

A fascinating article on "The Entropy of Black Holes," by V. F. Mukhanov, neatly ties the entropy of black holes, derived by Beckenstein and Hawking, to classical entropy from statistical mechanics. The entropy of a black hole is the logarithm of the number of ways it can be formed. Beautifully, the entropy of a black hole is proportional to its surface area, and the surface area is proportional to the square of the mass. Apparently, the black hole entropy is due to the loss of information about the structure of matter after this matter falls into a black hole. As Mukhanov says, "Because of the existence of an event horizon, it is impossible to obtain any information about the internal structure of such a black hole, and consequently we lose all information about the properties of the matter from which the black hole was formed." One of the interesting consequences of the entropy of the black hole is that if two black holes are coalesced, apparently by slowly lowering one black hole into another, the entropy does not double but quadruples.

One may be interested in the limits of the complexity of the universe. If, indeed, the universe is a computer, as argued by Fredkin, we might ask what its computational and memory capacities are. Here, Davies quotes the result of the Bekenstein-Hawking formula for black hole entropy to argue that if the entire universe were converted into a black hole, it would conceal a quantity of information I , given by

$$I \approx \frac{GM^2}{hc}$$

where M is the mass of the observable universe. Using current estimates, this results in $I = 10^{120}$. This limits the amount of blank tape available for a "universal" Turing machine.

There is really no question that physics and information theory are related. In fact, the word "entropy," used by Shannon in 1948 to designate the descriptive complexity of a random variable, owes itself to the work in statistical mechanics by Boltzmann, Clausius and others in the 19th century. It would be nice if the circle were completed, and the abstract information theory of Shannon were to play a role in physics again. However, the concept of physical information remains elusive.

I believe that there will eventually be such a theory. It will treat the number of distinguishable states of a physical system and the number of distinguishable states that can be reached from a given state. From this will evolve a physical theory of computation. We may even expect a second law of computation. Just as the second law of thermodynamics says that useful energy decreases, we can expect that useful information will decrease in a closed system. Thus, it may be impossible to build either a perpetual motion

machine or a perpetual computer. Physical systems dissipate into an informationally worthless homogeneous system of mass and energy.

In summary, this is a fascinating collection of papers on entropy and the physics of information. There are no final answers to the questions raised in this book. Nonetheless, the collection is accessible, has both elementary and advanced points of view on these questions, and is sure to be a vital stimulant in future thought on the subject.

Thomas M. Cover is a Professor jointly with the Departments of Electrical Engineering, and Statistics, at Stanford University, CA, working in the areas of communication theory, information theory, pattern recognition and complexity. He is coauthor with J. Thomas of the textbook, "Elements of Information Theory" (Wiley, New York), 1991. He was President of the IEEE Information Theory Society in 1972 and is a Fellow of the IMS, IEEE, and the AAAS.

Vector Quantization and Signal Compression, Allen Gersho and Robert M. Gray, (Boston: Kluwer Academic Publishers, 1992, 732 + xxii pages)

Reviewer: Nariman Farvardin, Senior Member, IEEE

In *Vector Quantization and Signal Compression*, Gersho and Gray provide a comprehensive and readable presentation of the theory and practice of signal compressions with an obvious emphasis on the recent developments on vector quantization. The book brings together, in a harmonious fashion, information and results from a multitude of developments on signal compression. Gersho and Gray have been two of the major contributors to these developments, which have taken place primarily since 1980, and are most qualified to author such a book.

The first book devoted to signal compression, *Digital Coding of Waveform*, was coauthored by Jayant and Noll in 1984. While the two books share the same objective, namely, analysis and design of efficient signal compression schemes, Gersho and Gray have i) emphasized the underlying principles and fundamental properties of signal compression schemes and ii) placed significantly more emphasis on vector quantization and its variants. The excellent book of Jayant and Noll is, for the most part, complementary to the book under review and should remain a useful reference in signal compression. In *Vector Quantization and Signal Compression*, Gersho and Gray expend their effort in presenting an algorithmic view of signal compression and how it might be applied to real-world compression paradigms. This is in contrast with the approach taken by R. M. Gray in *Source Coding Theory*, where he focuses on the information-theoretic aspects and fundamental performance limits of source coding as well as high rate quantization theory. This book can be used as a theoretical companion to the book under review.

Gersho and Gray have provided a balanced mix of mathematical rigor and engineering intuition to explain both the theoretical underpinnings and practical considerations of signal compression. In addition to being a useful reference for those already familiar with the basics of signal compression, the book can serve as a useful introduction for newcomers to the field. Currently, there is significant research interest in signal compression and there are reasons to believe that the trend will continue for some time. Thus, the publication of the book is very timely.

Apart from the introductory chapter, the book is divided into three parts. Part I, which contains three chapters, is intended to provide the

needed background and develop notation. The authors have carefully tailored their presentation to match the subsequent discussions in Parts II and III. In Part II a discussion of scalar quantization and other scalar quantizer-based systems is provided. Part III forms the core of the book and contains the most detailed discussion on vector quantization in the form of a book. In what follows a chapter-by-chapter review of the book is provided.

Chapter I provides an introduction to signal compression, including its history, significance and practical relevance. The notions of *lossless* and *lossy* compression are introduced and the main criteria for the design of a data compression system are discussed. In addition, practical considerations that rarely enter the mathematical optimization process, such as the encoding complexity and the choice of a perceptually meaningful distortion measure, are briefly discussed. Also, in this chapter an informative discussion on why signal compression is useful including a few examples of specific practical signal compression technologies, is provided.

Part I of the book which is comprised of Chapters 2–4, is intended to provide the tools needed for the understanding of the developments in the rest of the book. Part I serves two main purposes: 1) introducing the bulk of the notation used in the rest of the book and ii) rendering the book self-contained. Chapter 2 encapsulates the requisite knowledge of random processes and linear systems needed for an understanding of the subsequent chapters. First, a review of probability theory including random variables and vectors, random processes and expectation is provided. This is followed by a description of linear systems (both continuous- and discrete-time), an engineering and intuitive discussion of the notions of stationarity and ergodicity and, finally, specific examples of useful random processes.

In Chapter 3 the authors provide a fairly complete discussion of sampling. In addition to presenting the usual concepts such as periodic sampling, aliasing and imaging, more practical sampling issues such as additive sampling noise or sampling jitter are discussed. Two-dimensional sampling, including a general discussion of lattice sampling, is also presented. This chapter is useful in that it establishes the link between the discrete domain (time or space) in which the compression algorithms operate and the continuous domain in which most real-world signals (primarily speech and imagery) are defined.

Chapter 4 is entitled linear prediction. It begins by describing the basic principles of minimum mean squared error estimation and quickly focuses on linear estimation and the orthogonality principle. This, then leads to a presentation of finite-memory linear prediction, the development of Yule–Walker equations and the well-known Levinson–Durbin recursion. In addition, the problem of linear predictor design for empirical data where prior knowledge of autocorrelation coefficients is not available—a principal issue in most speech processing systems—is described and both autocorrelation and covariance methods are presented. Also, infinite memory linear prediction, the Wiener–Hopf equation and spectral factorization are discussed and a small but useful section on generating a discrete-time random process with a prescribed spectral density is included.

Part II, consisting of Chapters 5–9, is essentially devoted to scalar quantization and other coding schemes which use scalar quantization as the basic quantization operation, such as predictive and transform coding. By providing a complete treatment of scalar quantization, the authors place enough emphasis on what they call “the heart of analog-to-digital conversion” and at the same time pave the way for the discussion on vector quantization in Part III.