Order From Chaos

The random nature of random processes is uniquely determined by the information content of their structure

by Lynn Arthur Steen

One of the greatest achievements of 20th-century mathematics is the exploitation of random or stochastic models to bring intellectual order out of apparently chaotic natural processes. From statistical mechanics to human behavior, the hypothesis of randomness has made possible greater accuracy in predicting the course of events than did even the full elaboration of Newtonian mechanics.

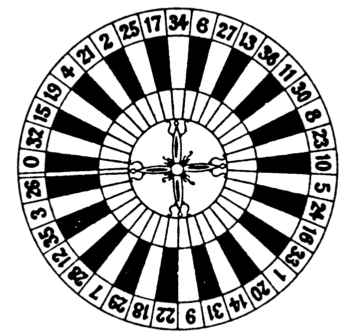
This shift away from pure determinism as an explanation of natural events has had major technical and intellectual repercussions, and still hides some significant unsolved problems on the interface of mathematics, science and philosophy.

In recent years several major results in theoretical probability have shed new light on many of these problems, offering not only better understanding of random processes but also new tools for future research. One of the most important advances is the proof that entropy provides a complete classification of independent random processes, as well as the discovery of a whole class of nondeterministic processes that cannot

"The consequences of these discoveries are likely to have major impact on the direction of future research on the nature of random processes."

be adequately modeled by traditional probabilistic structures. The consequences of these discoveries are likely to have major impact on the direction of future research on the nature of random processes, and, ultimately, on the nature of the mathematical structures employed in the modeling of natural phenomena.

Typically, mathematicians model random processes by clever manipulation of certain basic abstract processes such as the flipping of a coin, the spin of a roulette wheel or the draw of a card from a perfectly shuffled deck. They build their sophisticated stochastic models on a foundation abstracted from these simple physical processes because there is no compelling logical alterna-



"The shift away from pure determinism as an explanation of natural events has had major technical and intellectual repercussions, and still hides some significant unsolved problems on the interface of mathematics, science and philosophy."

tive. Although this procedure has succeeded well in developing nondeterministic models of particular processes, it has left unresolved certain major theoretical problems.

One such problem is whether there are any random processes that cannot be adequately modeled by clever conceptual use of a roulette wheel. The general roulette wheel with n possible outcomes, each with probability p_i (where $p_1 + p_2 + \ldots p_n = 1$) subsumes the coin flipping and card selection processes: Flipping a coin is tantamount to spinning a roulette wheel with two equally likely outcomes ($n = 2$, $p_1 = p_2 = 1/2$), while drawing a card from a shuffled deck is logically equivalent to spinning a roulette wheel with 52 equally likely slots. Such reasoning, usually in a more complex form, seems capable of reducing all nondeterministic models with a finite number of possible outcomes to a roulette model. What has not been known is whether the possibility of this reduction is necessarily so, or just accidentally so. Are there some processes, perhaps not yet discovered or studied, for which a roulette model, however intricate or complex, is totally inadequate?

A closely related question concerns the description and classification of the roulette models themselves. Probability theorists are intensely interested in these models bcause they are the building blocks for all higher models: Those who theorize must first understand these simple models in order to have any hope of clarifying their complex interaction in more sophisticated theories. It often happens that two different roulette models are only superficially different: They produce results with the same statistical structure, differing solely in external form. Such roulette models are called isomorphic (of the same form). Each can be coded into the other: The probabilistic information contained in either is equivalent to that in the other. Twenty years ago one of the major conjectures in this field was that all processes derived from a simple roulette model may be isomorphic to each other.

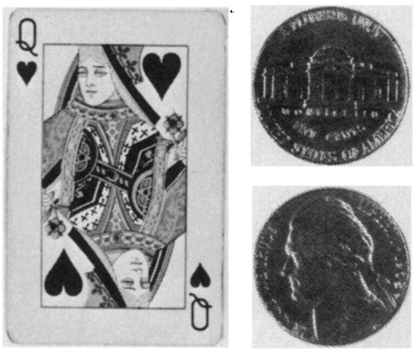
The first major development concerning this problem was the discovery around 1958 by the famous Russian probabilist A. N. Kolmogorov that the information contained in each roulette model was invariant under any legitimate coding. (Kolmogorov defined the

Are there any random processes that cannot be adequately modeled by clever conceptual use of a roulette wheel? The answer turns out to be: Yes.

information, or the entropy, of a roulette wheel, following Claude Shannon's work of a decade earlier, as $-\Sigma p_i \log p_i$). Kolmogorov's result implied that two roulette models with different entropy (or information) cannot be isomorphic to each other, for any translation of one into the other would leave the entropy unchanged (that is, invariant). Of course, there are infinitely many roulette models with the same entropy, and Kolmogorov's result does not say whether such models may or may not be isomorphic.

The study of these questions is part of the mathematical specialty known as ergodic theory. The principal tools of these investigations are certain transformations of the underlying probability space that preserve the probability measure. Among these so-called measure-preserving transformations are some, called Bernoulli shifts (after the famous 18th-century father of modern probability, Daniel Bernoulli) that are derived directly from roulette wheel models. The repeated spins of a roulette wheel may be represented by a string of symbols (stretching to infinity in both directions) each of which denote the outcome of a particular spin. The Bernoulli shift is the transformation that consists of translating (that is, shifting) this string of symbols one position to the right. Such a shift corresponds to a time translation of one unit.

Bernoulli shifts are the transformations in ergodic theory that correspond exactly to the independent processes of probability theory: The roulette wheel is the device that carries this correspondence. Bernoulli shifts arise in a variety of contexts, from multi-



step Markov processes to the Brownian motion of a hard-sphere gas. The fundamental question concerning the isomorphisms of the roulette process translate directly to an equivalent question concerning the isomorphisms of Bernoulli shifts. And this question has been recently settled by Stanford mathematician Donald Ornstein.

Ornstein has shown that Kolmogorov's entropy invariant provides a complete classification for Bernoulli shifts. Specifically, he proved that any two Bernoulli shifts with the same entropy are isomorphic. In other words, the information content of an independent process completely determines its probabilistic structure, except possibly for events with probability zero. Ornstein's result thus provides some insight into the epistemological status of random processes: Their random nature is uniquely determined by the information content (entropy) of their structure. Of course his theorem did much more than this on a technical level: It provided a simple and direct means of deciding whether two Bernoulli shifts (hence, any two independent random processes) are essentially the same.

Ornstein's work, which he has drawn together in a monograph, "Ergodic Theory, Randomness and Dynamical Systems," recently published by Yale University, leads to further insights into the nature of nondeterministic processes. Many such processes may be derived directly from Bernoulli shifts. These processes, called Bernoulli processes, are exactly those that can be approximated by finite coding of a roulette wheel—the longer the code, the better the approximation. They are in some sense the most random possible processes, and they are the only random processes that can be approximated well by a mechanism with a finite memory.

Ornstein and others have shown that any gross measurement—one with only a finite number of possible outcomes—on a mechanical system is a Bernoulli process: It produces a result essentially indistinguishable from a finite coding of a roulette wheel or a multistep Markov process. This result provides yet another clue concerning the relation between deterministic and nondeterministic phenomena: Gross measurement on a completely deterministic system yields the most random possible process!

A common desideratum of any random process is that behavior in the distant past should have little or no influence on the probabilities of present behavior. Kolmogorov proposed this as a criterion for completely nondeterministic processes. Specifically, he studied a class of processes, since called Kolmogorov processes, that satisfy the so-called "zero-one" law of probability theory: If knowledge of what the process did in the very distant past can help in any way to predict the present probability of a particular event, then that event must either have probability zero or one. A mechanical system has this property if and only if the only deterministic measurements that can be made on the system are those whose results are already certain in advance of the measurement. Every Bernoulli process is a Kolmogorov process, and Kolmogorov believed that the converse was also true, namely, every process that satisfies the zero-one law must have as its basic stochastic mechanism an independent process such as a roulette wheel.

Ornstein showed that this conjecture is false: He constructed an example of a completely nondeterministic process (that is, a process satisfying the zero-one law) that cannot be approximated by any multistep Markov process or by any finite coding of roulette wheels. This yields a third major insight into the nature of random processes: It is simply not true that all nondeterministic processes arise from a roulette-type mechanism. □

Lynn Arthur Steen is professor of mathematics at St. Olaf College, Northfield, Minn.

. . . Accelerators

much weaker than the electromagnetic if they are two aspects of the same thing. The mass of the boson reduces the strength of the weak interaction compared to the electromagnetic, whose carrier particle has zero mass.

In conclusion Weisskopf lists the next desirable steps in equipment and what they may hope to find in the light of these theoretical patterns.

First he proposes very high energy colliding beams of electrons and positrons (a positron is the antiparticle to an electron). The matter-antimatter annihilation that occurs when electron and positron meet produces a virtual photon, which then can turn itself into other particles. Such a collision, especially the production of what is called a timelike virtual photon, is a way of concentrating a large amount of energy into a small space. It is a good way of creating hadrons and studying their structure and of making previously unknown particles, as has lately been in the news. (SN: 11/23/74, p. 324) Because electron and positron have both electromagnetic and weak interactions, their collisions are also a good way to probe the unified theories.

A proton accelerator or proton-proton colliding beams that gave 100 billion electron-volts "in the center of mass"—thus made that much energy available for creation of new particles—could discover the intermediate vector bosons. It should be able to monitor high energy, and therefore short distance behavior of the strong interaction to see whether it really does go down. It might find free quarks if they can exist, and it might find exotic new particles. "If not," says Weisskopf, "the whole house of cards I have tried to build will collapse."

A fixed-target accelerator more energetic than the biggest now in existence could provide beams of secondary particles (neutrinos, pi mesons, K mesons, muons) with more than 200 billion electron-volts energy. These could test the unification of the weak and electromagnetic interactions, especially whether their strengths become equal in experiments where large amounts of momentum (comparable to the mass of an intermediate vector boson) are transferred from one particle to another. They could also test whether the strong interaction gets weak at high energies.

Weisskopf warns us to expect the unexpected—a good motto in particle physics. The theoretical patterns, ingenious though they are, are subject to correction by the phenomena, and they have a long history of that. "Very probably," he predicts, "all these ideas will turn out to be landing in India. People will discover a new continent, and this will be basic for our understanding of the structure of the universe." □