IS NATURE PROBABLE OR CAPRICIOUS?

Richard C. Lewontin

Probably the most important event in the history of modern science, although it is not always recognized as such, was the publication in 1637 of Descartes *Discourses*, especially of Part V, wherein he described the metaphor of the animal as a machine. Soon to be followed by La Méttrie's "*homme machine*," the "*bête machine*" of Descartes was the foundation of a world view that has colored science ever since and truly makes modern knowledge a Cartesian system. The concept of the universe as a machine is really the notion that for every cause there is an effect, that the universe is deterministic. Put in another way, it asserts that given the state of all the relevant variables at some time t_0, it is possible to predict exactly what the state of the universe will be at some future time t_1; that is, the universe can be described as a solution to a large set of simultaneous differential equations. This world view is responsible for vast progress in physics, chemistry, and physiological biology, and is, moreover, the starting point for work in any new science such as psychology. We always begin with the hope and assumption that rigorous laws leading to exact predictions are possible.

It is ironic that at the time of Descartes death in 1650 there was being laid the foundation of an alternate world view, one that would not have its full impact for more than 200 years. In the *De alea geometriae*, Pascal, Fermat, and Huyghens created an entirely new metaphor, that of the universe as a gaming table in which the cast of dice determined the outcome. The idea that human affairs were the outcome of chance events was, of course, a very old one, and Caesar at the Rubicon was not the first to say "*Alea jacta est*" at a critical moment. Nevertheless, we owe the explicit systemization of chance as an exact doctrine to Pascal and Fermat and its rigorous mathematization to Bernoulli, and much later to Laplace.

So great was the influence of Cartesianism on science, however, that even the founders of the probability theory were not willing to accept its antideterministic implications. Laplace in his *Essai philosophique sur les probabilitiés* reconciled chance events with a deterministic universe by the principle of ignorance. Events *seem* to be governed by chance because the immense complexity of the universe makes it impossible to know all the relevant facts about the state of the system. If there were a demon capable of comprehending all the important information about a die, including the exact forces involved in throwing it, its shape, the properties of the gaming table, and so on, then that demon could foretell exactly the outcome of a cast. For Laplace, then, the universe was still one of cause and effect, but human ignorance and fallibility made exact prediction impossible.

The first important physical applications of probabilistic notions came in the middle of the 19th century with the development by Maxwell and Boltzman of the kinetic theory of gases. But the influence of Laplace and Cartesianism was so great that even at the level of molecular events it was assumed that there was strict determination. Only our inability to distinguish among molecules made, in practice, a probabilistic system out of a deterministic one. There is a great similarity between Maxwell's demon and Laplace's.

It is only in the 20th century that physical science has accepted the full metaphysical implications of the notion of chance. Quantum mechanics is profoundly antimaterialistic and anti-Cartesian. It states, as a fundamental property of the universe, the probabilistic nature of events. No matter how much information there is about the past history of a given unstable nucleus, it is, in principle, impossible to decide when it will decay; that is, up until the instant of decay there is no difference between the nucleus that decays and its neighbor that does not. All that can be specified is the proportion of an ensemble that will decay in a given interval. Laplace and Maxwell postulate demons that do not exist in fact; Schroedinger and Heisenberg deny the possibility of their existence.

I have discussed the history of determinism and its eventual rejection as a universal because I would like to carry the process one step further and ask whether even the statements of a probabilistic universe are too certain for some phenomena. But this is a paradox. How can anything be less certain than chance? To understand this we must look at a fundamental axiom of probability theory, the Law of Large Numbers, and at the notion of statistical information.

The Law of Large Numbers states, in one of its many forms, that the average value of some variate will tend to lie in an interval around the true mean of that variate, which interval gets smaller and smaller as the number of observed values grows larger. As the sample size grows larger and larger, the observed mean is surer and surer to be closer and closer to the true value. This law holds for a very wide variety of cases, including mixtures of different distributions, and can be said to be the most fundamental axiom in probability theory. An example of its operation is shown in Figure 1. Fifty values were chosen from a table of random numbers and the mean of these numbers was taken using only the first number, then the first and second, then the first, second, and third, and so on up to a cumulative mean of all 50. These means are plotted against the sample size. The two solid lines represent the same group of random numbers taken in two different orders, while the dashed lines represent a completely different set of numbers also taken in two different orders. We see that for small sample sizes the means tend to deviate widely from the true mean of 5, but that as the samples grow larger and larger the means get closer and closer to the true value. Also evident is the tendency for the cumulative mean to remain on the same side of the true value for long periods. Finally, we see that the order in which the numbers are taken makes

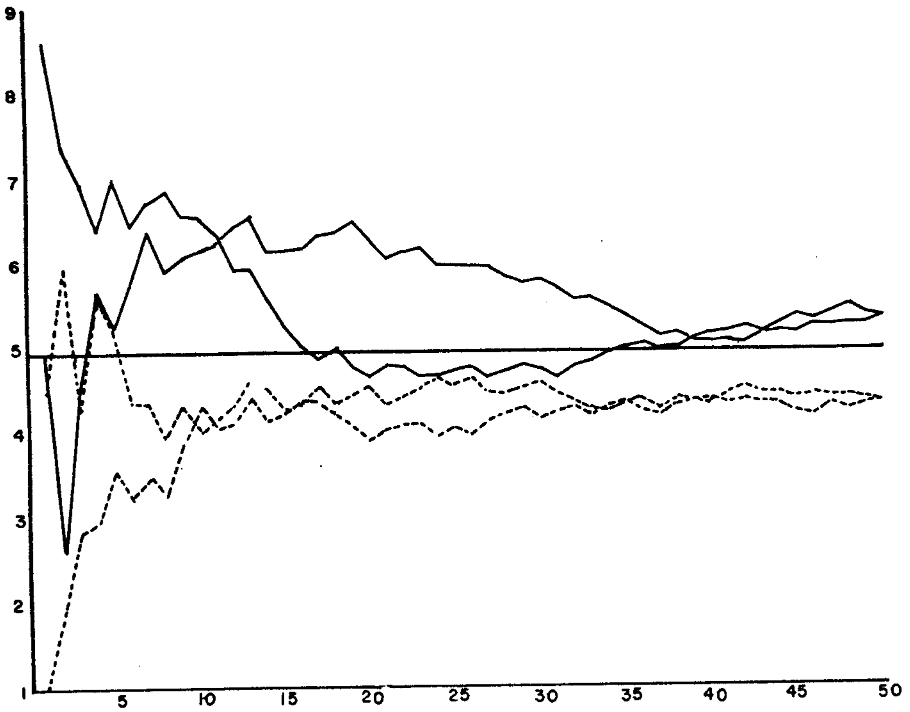

FIG. 1. Cumulative means different sequences of random numbers. Ordinate: cumulative mean; abscissa: length of sequence. The two solid lines are one set of random numbers in two different orders, the dashed lines are two different orders of a different set of numbers.

less and less difference as the sample size grows and, of course, when all 50 numbers are included, it makes no difference at all.

While this behavior of means (and of probabilities) is common over a great range of varieties, it is not universal. A famous case is that of the Cauchy distribution which looks superficially like a normal distribution but has somewhat higher probabilities of very large deviations. It is well-known that a mean based on a single observation from this distribution is just as close to the true mean as an average of a million observations. That is, if many random numbers were taken from a Cauchy distribution, Figure 1 would show wild jumps all along the abscissa with no tendency to approach the center. To put this in another way, a mean based on a single observation from the Cauchy distribution has just as much *information* about the true mean as a sample of size one million. The Law of Large Numbers might be restated as guaranteeing that the amount of information about the true mean contained in the sample mean grows larger and approaches perfect information as the sample grows larger.

Using this notion of information content of a sample, we can now go back to the question of determinism and chance. In a deterministic system there is perfect information about the true state of the universe (at least in theory) for any single occurrence of an event. In a probabilistic system, there is less than complete information in a finite set of observations, but the information approaches perfection as the size of the ensemble sample grows larger. I will call a system *capricious* if there is less than perfect information (often there is none) and if repeated sampling of the universe fails to increase the information about the system. We see then that capriciousness is a function of the mechanism for information storage available to the detector of the universe. The Law of Large Numbers is the law for a detector that *never forgets*, that is, as more and more data are fed into the mean, it is refined more and more and approaches the true value. But a detector that keeps forgetting may arrive at a steady state equilibrium between information input and information lost so that it cannot perfect itself. For such a detector the universe remains forever capricious because it is forever pulling unexpected tricks, repetitions of unremembered events.

Now organisms, looked at as detectors of the universe, have a limited memory and as individuals are subject to environmental caprice. But this is also true of populations and species through time, that is, evolutionary adaptation, looked at as acquiring information about the past environments in order to predict future ones, can never be perfect and therefore the environment, in some sense, will always be capricious. To illustrate the fact that the memory mechanism of populations is weaker than the Law of Large Numbers, I have used the random numbers of Figure 1 to construct an artificial case of evolution. We assume that a gene has two alleles, *a* and *A*, and that neither is dominant. We further assume that natural selection sometimes favors one allele and sometimes the other. In particular, the intensity of selection is equal to the deviation of our random numbers from their true mean of 5. There is a simple equation giving the frequency of the allele *a* in any generation from the frequency in the previous generation and the intensity of natural selection in that generation. Figure 2 shows the history of gene frequency changes in a population subject to a fluctuating selection. The selection intensities in successive generations are shown by the crosses. The solid line shows the history of gene frequency in a population subject to these selection pressures for 50 generations. The dashed line shows another population subject to the same selection pressures *but in reverse order*. We see that the two populations not only come to slightly different end points and have very different histories, but, more important, their *average behavior* has been totally different. The mean gene frequency of the solid line population over the 50 generations has been approximately 0.38, while it was 0.57 in the dashed line population. Moreover, the variance in gene frequency by time has been much greater in the dashed line population than in the solid line one.

The nature of gene frequency change is such that the gene frequency in a population does not have a simple informational relationship to past environments. If past environments have been mostly in one direction, the gene frequency will tend to extreme values (as in the solid line population near the middle of its life time) and when that happens it will be very insensitive to recent environments and will not respond to them. Thus, the low gene frequencies in the solid line population over the last half of its lifetime occur, despite the fact that nearly every value of selection (shown by the crosses) was above the mean during this period. While it is not obvious from the figure, the dashed line population generally contains more information about recent environments and less about past ones because its gene frequency remains near intermediate values.

In general, the dynamics of natural selection are such that gene frequencies contain little or no information about environments of the remote past, and are mostly a reflection of the not too distant past and of recent selection. This being the case, the genetic structure of a population is incapable of adjusting to environmental fluctuations that have a very long cycle time or a very short one, compared with generation length. Information about very short term fluctuations can be stored in physiological homeostatic devices. But environments that recur only with a long cycle time, no matter how frequent they have been, will represent a capricious nature.

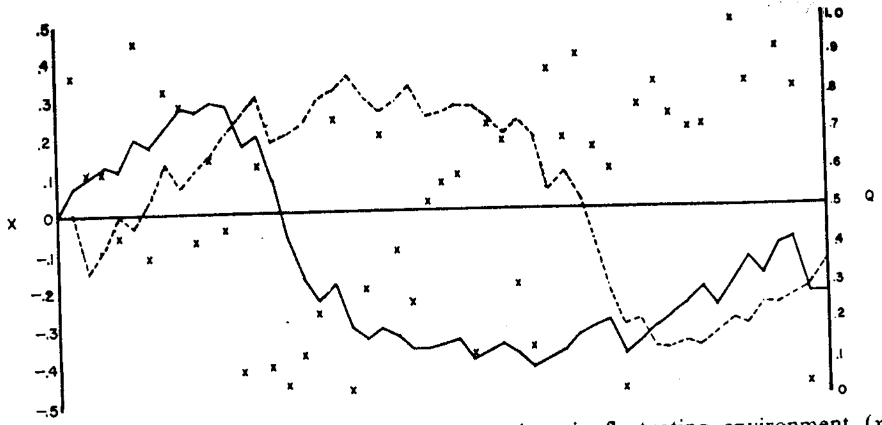

Fig. 2. Gene frequency (q) of successive generations in fluctuating environment (x). Crosses are environmental value. Solid line: gene frequency in successive generations caused by environments given by the crosses. Dashed line: gene frequencies resulting from same set of environments in reverse order.

Finally, Figure 2 illustrates the principal of *historicity* in evolution. The average genetic structure of a population in time depends not only on the static probability distribution of environments, *but on their historical sequence as well*. Even though environments may obey the Law of Large Numbers in terms of their means and other moments of their distributions, if the historical order in which environments occur is a significant variable in population adaptation, then an element of uniqueness is introduced. Again, for the population and the species, the accumulation of information is limited and nature is capricious.

LOGICAL TOOLS IN PHYSIOLOGY

William G. Van der Kloot

Whenever I read about the logical basis of science I become uncomfortable. I almost never feel like "an instrument for pointer readings in the hands of a disembodied intellect" (12); my laboratory life too often seems remote from the "delicate testing and remorseless logic" which Eddington (1) calls the soul of scientific progress. Many working scientists, I am sure, are unable to identify themselves and their daily operations with the automata specified by students of the "scientific method." Even an absolutely first-rate scientist will turn out on close acquaintance to be wonderfully human. "His ideas are not well ordered. He loves discussion but does not think always with c o m p l e t e l y consistent schemes, such as are used by philosophers, lawyers, or clergymen. Moreover in his laboratory he does not spend much of his time thinking about scientific laws at all" (15).

Many working physiologists are so nagged by misgivings about the rigor of their approach that they delight in stories belittling the part played by step-by-step logic in scientific creativity. For example, take the famous tale of Otto Loewi: picture him in his peaceful bed, Easter Eve, 1920 (6). In the middle of the night he dreams of an experiment perfectly designed to see whether the slowing of the heart produced by stimulating the vagus is brought about by a chemical released from the nerve ending. He immediately rouses himself, gropes on the bedside table for a pencil, and jots a brief note on what he always described — with 19th Century reticence — as a small square of thin paper.

In the morning he cannot read his own scrawl. Easter day is a torment, he recalls nothing of his idea, but, of course, on Easter evening, the dream is dreamed again. This time he goes straight to the laboratory and by day the work is done.

In evaluating this story we must first conclude that neither the setting nor the hour of the night make Loewi's achievement any less logical (if the life of the Professor then was like it is now, probably the day held no time for thought). The surprise is that he was suddenly so logical about this subject — he had not been working on junctional transmission at all. The action of the vagus in slowing the heart had been known for 50 years, and the *idea* of chemical transmission was not novel. What was new was Loewi's sudden vision of the resemblance between the idea of chemical transmission and the work going on in his own laboratory: he was studying the effects of digitalis on the heart of the frog. Digitalis was