

# Predicting with unpredictability

Random numbers: from stone casting to computers to radioactive decay, the generation of random sequences has always preoccupied mankind.

Gianpietro Malescio

**M**aking predictions is one of the main goals of science. Traditionally this implies writing down, and solving, the equations governing the system under investigation. When this method proves impossible we often turn to a stochastic approach. The term ‘stochastic’ encompasses a variety of techniques that are based on a common feature: using unpredictable entities — random numbers — to make predictions possible.

The origins of stochastic simulation can be traced to an experiment performed in the eighteenth century by Georges Louis Leclerc, Comte de Buffon. Leclerc repeatedly tossed a needle at random on to a board ruled with parallel straight lines. From his observations, he derived the probability that the needle would intersect a line. Subsequently, Pierre Simon de Laplace saw in this experiment a way to obtain a statistical estimate of pi.

Later, the advent of mechanical calculating machines allowed numerical ‘experiments’ such as that performed in 1901 by William Thomson (Lord Kelvin) to demonstrate the equipartition theorem of the internal energy of a gas. Enrico Fermi was probably the first to apply statistical sampling to research problems, while studying neutron diffusion in the early 1930s. During their work on the Manhattan Project, Stanislaw Ulam, John von Neumann and Nicholas Metropolis rediscovered Fermi’s method. They established the use of random numbers as a formal methodology, generating the ‘Monte Carlo method’ — named after the city famous for its gambling facilities. Today, stochastic simulation is used to study a wide variety of problems (many of which are not at all probabilistic), ranging from the economy to medicine, and from traffic flow to biochemistry or the physics of matter.

When the temporal evolution of a system cannot be studied by traditional means, random numbers can be used to generate an ‘alternative’ evolution. Starting with a possible configuration, small, random changes are introduced to generate a new arrangement: whenever this is more stable than the previous one, it replaces it, usually until the most stable configuration is reached. Randomness cannot tell us where the system likes to go, but allows the next best thing: exploration of the space of the configurations while avoiding any bias that might exclude the region of the



Predicting the unpredictable: casting rune stones has been used as a divination tool for more than 2,000 years.

possible solution. If we are able to guess the probability distribution of the configurations, then instead of conducting a uniform random search we can perform an ‘importance’ sampling, focusing our search on where the solution is more likely to be found.

Optimization problems are often solved using stochastic algorithms that mimic biological evolution. Although it may sound vaguely unpleasant, we come from a random search. In nature, new genetic variants are introduced through random changes (mutations) in the genetic pool while additional variability is provided by the random mixing of parent genes (by recombination). Randomness allows organisms to explore new ‘designs’ which the environment checks for fitness, selecting those most suited to survival. But the optimal solution is not found once and for ever. A continually changing environment means evolution is an ongoing process; it does not produce the ‘perfect’ organism, but rather a dynamic balance of myriad organisms within an ecosystem.

Generating true randomness is a challenging task. Early attempts at stochastic simulation produced samples through processes such as dice tosses or card draws. These phenomena in principle obey newtonian mechanics, but in practice evolve unpredictably owing to their chaotic dynamics. Computers have made it much easier to produce large numbers of samples. They cannot, however, generate true random numbers.

As Gregory Chaitin and Andrei Nikolaevich Kolmogorov first pointed out in the 1960s, a random sequence is algorithmically incompressible; that is, it cannot be represented with a description shorter than the sequence itself. A string of ‘random’ numbers generated by a computer can be easily compressed since it can be described by the

algorithm that generates the sequence. Although computers can produce pseudorandom numbers that are virtually indistinguishable from true random numbers, all such sequences eventually show correlations, and fall into a repeating pattern.

To circumvent these problems, scientists are studying physical devices that can produce efficiently random numbers. A variety of phenomena have been considered, including the detection of atmospheric chaos by a noisy radio receiver, the timing of radioactive decays, and the beaming of photons at a semi-transparent mirror. In principle, random numbers captured in the real

world are superior to those generated by computers. In practice, sequences of pseudorandom numbers may seem more convincingly random than those produced by hardware generators.

Devising random-number generating physical processes that are precisely balanced is not easy. For example, coin tossing — a seeming paradigm of randomness — is actually subtly biased owing to the mass imbalance derived from the different designs on the two sides of the coin. Even if carefully engineered to minimize bias and other artefacts, hardware generators need ‘whitening’ algorithms to treat residual correlations, and statistical tests to detect possible failures in randomness. This can be checked only by using computers, which are notoriously unable to produce true random numbers! Thus, more reliable and efficient generators of random numbers cannot rely solely on either hardware or software, but should employ a combination of both.

Creating perfect disorder, eliminating even the slightest bias or correlation, is as difficult a task as maintaining perfect order. Nevertheless, the search for pure noise will go on because the usefulness of random numbers for making predictions resides in their complete unpredictability. As Chaitin puts it: “The world consists of the tension between order and chaos. When simulating physical phenomena, order is supplied by the laws of physics and chaos is supplied by random numbers.”

Gianpietro Malescio is in the Department of Physics, University of Messina, 98166 Messina, Italy.

## FURTHER READING

Metropolis, N. *Los Alamos Sci.* **15**, 125–130 (1987).

Chaitin, G. J. *Sci. Am.* **232**, 47–52 (1975).

Calude, C. S. & Chaitin, G. J. *Nature* **400**, 319–320 (1999).