What is Complexity?

Remarks on simplicity and complexity
by the Nobel Prize-winning author of The Quark and the Jaguar

MURRAY GELL-MANN

Nobel laureate Murray Gell-Mann has many intellectual passions including natural history, linguistics, archeology, history, depth psychology and creative thinking. As Professor and Co-Chairman of the Science Board of the Santa Fe Institute, he works with many researchers in their search for connections between basic laws of physics and the complexity and diversity of the natural world.

What is complexity? A great many quantities have been proposed as measures of something like complexity. In fact, a variety of different measures would be required to capture all our intuitive ideas about what is meant by complexity and by its opposite, simplicity.

Some of the quantities, like computational complexity, are time (or space) measures. They are concerned with how long it would take (or how much capacity would be needed), at a minimum, for a standard universal computer to perform a particular task. Computational complexity itself is related to the least time (or number of steps) needed to carry out a certain computation.

Other suggested quantities are information measures, referring, roughly speaking, to the length of the shortest message conveying certain information. For example, the algorithmic information content (or AIC) of a string of bits is defined as the length of the shortest program that will cause a standard universal computer to print out the string of bits and then halt.

As measures of something like complexity for an entity in the real world, all such quantities are to some extent context-dependent or even subjective. They depend on the coarse graining (level of detail) of the description of the entity, on the previous knowledge and understanding of the world that is assumed, on the language employed, on the coding method used for conversion from that language into a string of bits, and on the particular idealized computer chosen as a standard. However, if one is considering a sequence of similar entities of increasing size and complexity, and one is interested only in how the measure behaves as the size becomes large, then of course many of the arbitrary features become negligible. Students of computational complexity are typically concerned with whether a sequence of larger and larger problems can be solved in a time that grows as a polynomial in the size of the problem (rather than an exponential or something worse). It is probably safe to say that any measure of complexity is most useful for comparisons between things at least one of which has high complexity by that measure.

Many of the candidate quantities are uncomputable. For example, the algorithmic information content of a long bit string can readily be shown to be less than or equal to some value. But for any such value there is no way of excluding the possibility that the AIC could be lower, reduced by some as yet undiscovered theorem revealing a hidden regularity in the string. A bit string that is incompressible has no such regularities and is defined as "random." A random bit string has maximal AIC for its length, since the shortest program that will cause the standard computer to print it out and then halt is just the one that says PRINT followed by the string itself.

This property of AIC, which leads to its being called, on occasion, "algorithmic randomness," reveals the unsuitability of the quantity as a measure of complexity, since

the works of Shakespeare have a lower AIC than random gibberish of the same length that would typically be typed by the proverbial roomful of monkeys.

A measure that corresponds much better to what is usually meant by complexity in ordinary conversation, as well as in scientific discourse, refers not to the length of the most concise description of an entity (which is roughly what AIC is), but to the length of a concise description of a set of the entity's regularities. Thus something almost entirely random, with practically no regularities, would have effective complexity near zero. So would something completely regular, such as a bit string consisting entirely of zeroes. Effective complexity can be high only in a region intermediate between total order and complete disorder.

There can exist no procedure for finding all regularities of an entity. But classes of regularities can be identified. Finding regularities typically refers to taking the available data about the entity, processing them in some manner into, say, a bit string, and then dividing that string into parts in a particular way and looking for mutual AIC among the parts. If a string is divided into two parts, for example, the mutual AIC can be taken to be the sum of the AIC's of the parts minus the AIC of the whole. An amount of mutual algorithmic information content above a certain threshold can be considered diagnostic of a regularity. Given the identified regularities, the corresponding effective complexity is the AIC of a description of those regularities.

More precisely, any particular regularities may be regarded as embedding the entity in question in a set of entities sharing the regularities and differing only in other respects. In general, the regularities associate a probability with each entity in the set. (The probabilities are in many cases all equal but they may differ from one member of the set to another.) The effective complexity of the regularities can then be defined as the AIC of the description of the set of entities and their probabilities. (Specifying a given entity, such as the original one, requires additional information.)

Some authors have tried to characterize complexity by using the *amount* of mutual algorithmic information rather than the length of a concise description of the corresponding regularities. Such a choice of measure does not agree very well, however, with what is usually meant by complexity. Take, as a simple example, any string of bits consisting entirely of pairs 00 and 11. Such a string possesses an obvious regularity, but one that can be very briefly described: the sequences of odd-numbered and even-numbered bits are identical. The quantity of mutual AIC between those sequences is enormous, however, for a long string. Evidently the complexity here is better represented by the length of the brief description than by the amount of mutual algorithmic information.

Since it is impossible to find all regularities of an entity, the question arises as to who or what determines the class of regularities to be identified. One answer is to point to a most important set of systems, each of which functions precisely by identifying certain regularities in the data stream reaching it and compressing those regularities into a concise package of information. The data stream includes information about the system itself, its environment, and the interaction between the environment and the behav-

ior of the system. The package of information or "schema" is subject to variation, in such a way that there is competition among different schemata. Each schema can be used, along with some of the data, to describe the system and its environment, to predict the future, and to prescribe behavior for the system. But the description and prediction can be checked against further data, with the comparison feeding back to influence the competition among schemata. Likewise behavior conforming to a prescription has real world consequences, which can also affect the competition. In this way the schemata evolve, with a general tendency to favor better description and prediction as well as behavior conforming more or less to the selection pressures in the real world.

Examples on Earth of the operation of complex adaptive systems include biological evolution, learning and thinking in animals (including people), the functioning of the immune system in mammals and other vertebrates, the operation of the human scientific enterprise, and the behavior of computers that are built or programmed to evolve strategies—for example by means of neural nets or genetic algorithms. Clearly, complex adaptive systems have a tendency to give rise to other complex adaptive systems.

It is worth remarking for readers of this journal that John Holland, for example, uses a different set of terms to describe some of the same ideas. He uses "adaptive agent" for a complex adaptive system as defined above, reserving the name "complex adaptive system" for a composite complex adaptive system (such as an economy or an ecological system) consisting of many adaptive agents making predictions of one another's

> **Effective complexity can be high only in a region intermediate between total order and complete disorder.**

ESSAYS & COMMENTARIES

behavior. What I call a schema he calls an internal model. Both of us are conforming to the old saying that a scientist would rather use someone else's toothbrush than another scientist's nomenclature.

Any complex adaptive system can, of course, make mistakes in spotting regularities. We human beings, who are prone to superstition and often engage in denial of the obvious, are all too familiar with such errors.

Besides the possibility of error, we should also consider the difficulty of computation. How much time is involved in deducing practical predictions from a highly compressed schema, say a scientific theory, together with some specific additional data such as boundary conditions? Here we encounter time measures of "complexity," for instance logical depth, which for a bit string is related to the time required for a standard universal computer to compute the string, print it out, and then halt. That time is averaged over the various programs that will accomplish the task, with an averaging procedure that weights shorter programs more heavily. We can then consider the logical depth of any entity if a suitably coarse-grained description of it is encoded into a bit string.

A kind of inverse concept to logical depth is crypticity, which measures the time needed for a computer to reverse the process and go from a bit string to one of the shorter programs that will generate it. In the human scientific enterprise, we can identify crypticity roughly with the difficulty of constructing a good theory from a set of data, while logical depth is a crude measure of the difficulty of making predictions from the theory.

It is often hard to tell whether something that is apparently complex really possesses a great deal of effective complexity or reflects instead underlying simplicity combined with a certain amount of logical depth. Faced with a fairly detailed diagram of Mandelbrot's famous fractal set, for example, we might attribute to it a high effective complexity until we learn that it can be generated from a very simple formula. It has logical depth (and not even a gigantic amount

of that) rather than effective complexity. In contemplating natural phenomena, we frequently have to distinguish between effective complexity and logical depth. For example, the apparently complicated pattern of energy levels of atomic nuclei might easily be misattributed to some complex law at the fundamental level, but it is now believed to follow from a simple underlying theory of quarks, gluons, and photons, although lengthy calculations would be required to deduce the detailed pattern from the basic equations. Thus the pattern has a good deal of logical depth and very little effective complexity.

> **In contemplating natural phenomena, we frequently have to distinguish between effective complexity and logical depth.**

It now seems likely that the fundamental law governing the behavior of all matter in the universe—the unified quantum field theory of all the elementary particles and their interactions—is quite simple. (In fact, we already have a plausible candidate in the form of superstring theory.) It also appears that the boundary condition specifying the initial condition of the universe around the beginning of its expansion may be simple as well. If both of these propositions are true, does that mean that there is hardly any effective complexity in the universe? Not at all, because of the relentless operation of chance.

Given the basic law and the initial condition, the history of the universe is by no means determined, because the law is quantum-mechanical, thus yielding only probabilities for alternative histories. Moreover, histories can be assigned probabilities only if they are sufficiently coarse-grained to display decoherence (the absence of interference terms between them). Thus quantum mechanics introduces a great deal of indeterminacy, going far beyond the rather trivial indeterminacy associated with

Heisenberg's uncertainty principle.

Of course, in many cases the quantum-mechanical probabilities are very close to certainties, so that deterministic classical physics is a good approximation. But even in the classical limit, and even when the laws and initial condition are exactly specified, indeterminacy can still be introduced by any ignorance of previous history. Moreover, the effects of such ignorance can be magnified by the phenomenon of chaos in nonlinear dynamics, whereby future outcomes are arbitrarily sensitive to tiny changes in present conditions.

We can think of the alternative possible coarse-grained histories of the universe as forming a branching tree, with probabilities at each branching. Note these are *a priori* probabilities rather than statistical ones, unless we engage in the exercise of treating the universe as one of a huge set of alternative universes, forming a "multiverse." Of course, even within a single universe cases arise of reproducible events (such as physics experiments), and for those events the *a priori* probabilities of the quantum mechanics of the universe yield conventional statistical probabilities.

Any entity in the world around us, such as an individual human being, owes its existence not only to the simple fundamental law of physics and the boundary condition on the early universe but also to the outcomes of an inconceivably long sequence of probabilistic events, each of which could have turned out differently.

Now a great many of those accidents, for instance most cases of the bouncing of a particular molecule in a gas to the right rather than the left in a molecular collision, have few ramifications for the future coarse-grained histories. Sometimes, however, an accident can have widespread consequences for the future, although those are typically restricted to particular regions of space and time. Such a "frozen accident" produces a great deal of mutual algorithmic information among various parts or aspects of a future coarse-grained history of the universe, for many such histories and for various ways of dividing them up.

But such a situation, in which there is

a great deal of mutual algorithmic information generated, corresponds precisely to what we have called a regularity. Thus, as time goes by in the history of the universe and accidents (with probabilities for various outcomes) accumulate, so do frozen accidents, giving rise to regularities. Most of the effective complexity of the universe lies in the AIC of a description of those frozen accidents and their consequences, while only a small part comes from the simple fundamental laws of the universe, (the law of the elementary particles and the condition at the beginning of the expansion). For a given entity in the universe, it is of course only the frozen accidents leading up to its own regularities that contribute, along with the basic laws, to its effective complexity.

As the universe grows older and frozen accidents pile up, the opportunities for effective complexity to increase keep accumulating as well. Hence there is a tendency for the envelope of complexity to expand, even though any given entity may either increase or decrease its complexity during a given time period.

The appearance of more and more complex forms is not a phenomenon restricted to the evolution of complex adaptive systems, although for those systems

> **As the universe grows older and frozen accidents pile up, the opportunities for effective complexity to increase keep accumulating as well.**

the possibility arises of a selective advantage being associated under certain circumstances with increased complexity.

The second law of thermodynamics, which requires average entropy (or disorder) to increase, does not in any way forbid local order from arising through various mechanisms of self-organization, which can turn accidents into frozen ones producing extensive regularities. Again, such mechanisms are not restricted to complex adaptive systems.

Different entities may have different potentialities for developing higher complexity. Something that is not particularly distinguished from similar things by its effective complexity can nevertheless be remarkable for the complexity it may achieve in the future. Therefore it is important to define a new quantity, "potential complexity," as a function of future

time, relative to a fixed time, say the present. The new quantity is the effective complexity of the entity at a future time, averaged over the various coarse-grained histories of the universe between the present and that time, weighted according to their probabilities.

The era may not last forever in which more and more complex forms appear as time goes on. If, in the very distant future, virtually all nuclei in the universe decay into electrons and positrons, neutrinos and antineutrinos, and photons, then the era characterized by fairly well-defined individual objects may draw to an end, while self-organization becomes rare and the envelope of complexity begins to shrink.

ACKNOWLEDGMENTS

Technological Evolutions and Adaptive Organizations
by Stuart Kauffman and William Macready

Models and Aphorisms
by David Lane

Forthcoming Articles

A Language of Structure in the Science of Complex Systems
by Jeffrey Johnson