2

Complexity in the Universe

Charles H. Bennett
IBM Research Division, T. J. Watson Research Center
Yorktown Heights, NY 10598, USA.

Abstract

Nontrivial, "complex" or "organized" states of a physical system may be characterized as those implausible save as the result of a long causal history or evolution. This notion, formalized by the tools of the theory of universal digital computers, is compared to other notions of complexity, and an attempt is made to sketch open problems in the computation theory and statistical physics whose resolution would lead to a better fundamental understanding of "self-organization" in the universe.

The manifest complexity of many parts of the universe, especially living organisms and their byproducts, was formerly thought to be an expression of divine creativity, but is now widely believed to result from a general capacity of matter, implicit in known physical laws, to "self-organize" under certain conditions.

As a rough illustration of the essential ideas of self-organization, consider a sealed aquarium illuminated by a light source and containing a dead mixture of prebiotic chemicals. After a long time this system will fluctuate or otherwise find its way into a "live" macrostate containing, for example, fish and green plants. This live state will be somewhat stabilized relative to dead states by the light, which enables the organisms to grow, reproduce and metabolically defend themselves against thermal and other degradation. In an aquarium of ordinary size, dead states would still be overwhelmingly more probable even in the presence of the light, because spontaneous biogenesis would probably be far less likely than a "gambler's ruin" ecological fluctuation in which, say, the fish died of starvation after eating all the plants. But the larger the aquarium, the less likely will be a simultaneous extinction everywhere within it, and the more likely a spontaneous biogenesis somewhere within it. Finally, if the aquarium were the size of the earth, it might spend most of its time alive, as suggested by paleontological evidence of one dead-to-live and no live-to-dead transitions so far.

To elevate this kind of thinking from a truism that everyone agrees with but no one really understands, to a level of provable or refutable conjectures in statistical physics, we need a more rigorous and mathematical definition of "complexity," the quantity that supposedly increases when a self-organizing system organizes itself. As might be expected, the problem of defining complexity is itself complex, and there are many satisfactory definitions of different kinds of complexity. Below we compare a number of candidates for a definition of complexity, dismissing most of them as unsuitable to our purpose, without meaning to disparage their appropriateness in other contexts. For further details see [2][4][5].

An object might be considered complex if it has *complex behavior or function*, for example if it is able to grow, reproduce, adapt, or evolve in an appropriate environment. Even if it were possible to find mathematical definitions of these properties, we believe a more structural and less functional definition is needed to understand self-organization, because even functionally inert objects, such as a dead body, a book, or a fossil, subjectively can be said to be complex, and would not plausibly be found in a universe lacking some sort of self-organization, or, God forbid, divine intervention.

A more mathematical property related to complex function is *computational universality*, the ability of a system to be programmed through its initial condition to simulate any digital computation. Originally demonstrated for computer-like models such as Turing machines and deterministic cellular automata, computational universality has subsequently been demonstrated for models more closely resembling those studied in mechanics and statistical mechanics, e.g. the hard sphere gas in an appropriate periodic potential [10], noisy cellular automata in 1 and 3 dimensions [12] [13], systems of partial differential equations [17] and even a single classical particle in a finitely complicated box [16]. The ability of universal systems to simulate one another entails that the dynamics of any one of them encodes, in a straightforward manner, the dynamics of any other, and indeed of any process whose outcome can be determined by logical deduction or numerical simulation. For example, one can readily find an initial condition for Moore's particle [16] which will enter a designated region of space if and only if white has a winning strategy in chess, and another initial condition that will do so if and only if the millionth decimal digit of π is a 7. Computational universality therefore now appears to be a property that realistic physical systems can have; moreover if a physical system does have that property, it is by definition capable of behavior as complex as any that can be digitally simulated.

However, computational universality is an unsuitable criterion of complexity for our purposes because it is a functional property of systems rather than a structural property of states. In other words it does not distinguish between a system merely capable of complex behavior and one in which the complex behavior has actually occurred. The complexity measure we will ultimately advocate, logical depth, is closely related to the notion of universal computation, but it allows complexity to increase as it intuitively should in the course of a "self-organizing" system's time development.

Thermodynamic potentials, such as entropy or free energy, measure capacity for irreversible change, but do not agree with subjective complexity. A human body is more complex than a vat of nitroglycerine, but has lower free energy. Similarly a bottle of sterile nutrient solution has higher free energy, but lower subjective complexity, than the bacterial culture it would turn into if inocculated with a single seed bacterium. The growth of bacteria following inocculation is a thermodynamically irreversible process analogous to crystallization of a supersaturated solution inocculated with a seed crystal. Each is accompanied by a decrease in free energy, and, even in the absence of a seed, is vastly more probable than its reverse: the spontaneous melting of a crystal into a supersaturated solution, or the spontaneous transformation of bacteria into high-free-energy nutrients. The unliklihood of a bottle of sterile nutrient transforming itself into bacteria is therefore not a manifestation of the second law, but rather of a putative new "slow growth law" which forbids complexity, however it is defined, to increase quickly, but allows it to increase slowly, e.g. over geological time in biogenesis. This example also illustrates the non-additivity of subjective complexity. One bacterium seems much more complex than none, but only slightly less complex than the bottle full of descendants it can quickly give rise to.

Algorithmic Information Content, also called Algorithmic Entropy, Algorithmic Complexity, or Solomonoff-Kolmogorov-Chaitin Complexity [20][7][8], formalizes the notion of amount of information necessary to uniquely describe a digital object x. A digital object means one that can be represented as a finite binary string, for example, a genome, an Ising microstate, or an appropriately coarse-grained representation of a point in some continuum state space. The algorithmic entropy $H(x)$ of such an object is defined as the negative base-2 logarithm of the object's *algorithmic probability*, $P(x)$. This in turn is defined as the probability that a standard universal computer U, randomly programmed (for example by the proverbial monkey typing at a binary keyboard with two keys), would embark on a computation yielding x as its sole output, afterward halting. The algorithmic probability $P(x)$ may be thought of a weighted sum of contributions from all programs that produce x, each weighted according to the negative exponential of its binary length, which is the probability that the monkey will type that particular program and so cause it to be executed. An *algorithmically random* string is defined as one of maximal information content, nearly equal to the length of the string (even if a string has no regularities permitting it to be produced with higher probability, any N–bit string can be generated with probability at least $2^{-(N+O(\log N))}$ by a "print program" in which the monkey essentially types the string out verbatim, along with instructions, of length $O(\log N)$, directing the computer to pass these N bits on directly to the output and then halt).

Turning now to the sum of $P(x)$ over *outputs*, this sum $\sum_x P(x)$ is not equal to unity as one might first suppose, because, as is well known, an undecidable subset of all universal computations fail to halt, and so produce no output. Therefore $\sum_x P(x)$ is an uncomputable irrational number less than 1. This number, called

Chaitin's Omega [7], has many remarkable properties [14], such as the fact that its uncomputable digit sequence is a maximally compressed form of the information required to *solve* the halting problem.

Despite being defined in terms of a particular universal computer, algorithmic probability is machine-independent up to an multiplicative constant (and algorithmic entropy up to an additive constant), because of the ability of universal computers to simulate one another (programs for one machine can be adapted to run on another by prefixing each program with a constant string, directing the second machine to simulate the first).

Though very differently defined, algorithmic entropy is typically very close to ordinary statistical entropy $-\sum p \log p$ in value. To take a simple example, it is easy to show that almost all N–bit strings drawn from a uniform distribution (of statistical entropy N bits) have algorithmic entropy nearly N bits. More generally, in any concisely describable ensemble of digital objects, e.g. a canonical ensemble of Ising microstates at a given temperature, the ensemble average of the objects' algorithmic entropy closely approximates the whole ensemble's statistical entropy [20] [1]. In the case of continuous ensembles, the relation between algorithmic and statistical entropy is less direct because it depends on the choice of coarse-graining. Zurek [21] discusses some of the conceptual issues involved.

For this reason algorithmic information is best thought of as a measure of randomness, not subjective complexity, being maximal for coin-toss sequences, which are among the least organized subjectively. Typical organized objects, on the other hand, precisely because they are partially constrained and determined by the need to encode coherent function or meaning, contains less information than random sequences of the same length; and this information reflects not their organization, but their residual randomness.

For example, the algorithmic information content of a genome represents the extent to which it is underdetermined by the constraint of viability. The existence of noncoding DNA, and the several percent differences between proteins performing apparently identical functions in different species, make it clear that a sizable fraction of the genetic coding capacity is given over to transmitting such "frozen accidents", evolutionary choices that might just as well have been made otherwise.

A better way of applying statistical or algorithmic information to the definition of organization is to use it to characterize the correlations typical of organized or complex objects: two parts of such an object taken together typically require fewer bits to describe than the same two parts taken separately. This difference, the *mutual algorithmic information* between the parts, is the algorithmic counterpart of the non-additivity of statistical or thermodynamic entropy between the two parts, the amount by which the entropy of the whole falls short of the sum of the entropies of the two parts. In many contexts, e.g., communications through a noisy channel, mutual information can be viewed as the "meaningful" part of a message's information, the rest being meaningless information or "noise".

A body is said to have long range order if even arbitrarily remote parts of it are correlated. However, crystals have long range order but are not subjectively very complex. Organization has more to do with the *amount* of long-range correlation, i.e., the number of bits of mutual information between remote parts of the body. Although we will ultimately recommend a different organization measure (logical depth), remote mutual information merits some discussion, because it is characteristically formed by nonequilibrium processes, and can apparently be present only in small amounts at thermal equilibrium.

If two cells are taken from opposite ends of a multicellular organism, they will have a large amount of mutual information, if for no other reason than the presence in each cell of the same genome with the same load of frozen accidents. As indicated earlier, it is reasonably certain that at least several per cent of the coding capacity of natural genomes is used to transmit frozen accidents, and hence that the mutual information between parts of a higher organism is at least in the hundred megabit range. More generally, mutual information exists between remote parts of an organism (or a genome, or a book) because the parts contain evidence of a common, somewhat accidental history, and because they must function together in a way that imposes correlations between the parts without strictly determining the structure of any one part. An attractive feature of remote mutual information for physical systems is that it tends to a finite limit as the fineness of coarse-graining is increased, unlike simple information or entropy in a classical system.

Since mutual information arises when an accident occurring in one place is replicated or propagated to another remote place, its creation is an almost unavoidable side effect of reproduction in a probabilistic environment. Another obvious connection between mutual information and biology is the growth of mutual information between an organism and its environment when the organism adapts or learns.

Further support for remote mutual information as an organization measure comes from the fact that systems stable at thermal equilibrium, even those with long range order, exhibit much less of it than nonequilibrium systems. Correlations in systems at equilibrium are generally of two kinds: short range correlations involving a large number of bits of information (e.g. the frozen-in correlations between adjacent lattice planes of an ice crystal, or the instantaneous correlations between atomic positions in adjacent regions of any solid or liquid), and long range correlations involving only a few bits of information. Typical of these latter correlations are infinite-range correlations associated with order parameters such as magnetization and crystal lattice orientation and phase. Even when these order parameters are continuous, they convey only a few bits of information, owing to the thermal and zero-point disorder which causes the lattice orientation, say, of an N–atom crystal to be well-defined only to about $\log N$ bits precision. Besides involving much less information, remote correlations at equilibrium differ qualitatively from the non-equilibrium ones discussed earlier: equilibrium correlations, in a system with short-range forces, must be propagated through an intervening medium, while nonequililbrium ones (e.g.

between the contents of two newspaper dispensers in the same city) need not pass through the intervening medium but are instead typically propagated through a V-shaped path in spacetime connecting the random origin of the information at an earlier time with two separated copies of it at a later time.

Despite these advantages, we believe remote mutual information is an unsatisfactory complexity measure because large quantities of it can be produced rapidly, by subjectively trivial nonequilibrium processes, in violation of the slow growth law. For example, by pulverizing a piece of glass with a hammer, one can produce a kind of 3-dimensional jigsaw puzzle of atomically complementary random fracture surfaces, with a non-additivity of entropy, between two specimens of the powder, proportional to the area of complementary surface between them. A greater non-additivity could be produced by enzymatically replicating, and then stirring, a solution of random, biologically meaningless DNA molecules to produce a kind of jigsaw puzzle soup, two spoonfuls of which would have macroscopically less than twice the entropy of one spoonful. In both these examples, the mutual information is formed by nonequilibrium processes and would decay if the system were allowed to approach a state of true thermal equilibrium, e.g. by annealing of the separated fracture surfaces.

A conspicuous feature of many nontrivial objects in nature and mathematics is the possession of a *fractal or self-similar or hierarchical structure,* in which a part of the object is identical to, or is described by the same statistics as, an appropriately scaled image of the whole. This often beautiful property is too specialized to be an intuitively satistfactory criterion of complexity because it is absent from some subjectively complex objects, such as the decimal expansion of pi, and because, on the other hand self-similar structures can be produced quickly, e.g. by deterministic cellular automata, in violation of the slow growth law. Even so, the frequent association of self-similarity with other forms of organization deserves comment. In some cases, self-similarity is a side-effect of computational universality, because a universal computer's ability to simulate other computers gives it in particular the ability to simulate itself. This makes the behavior of the computer on a subset of its input space (e.g., all inputs beginning with some prefix s that tells the computer to simulate itself) replicate its behavior on the whole input space.

Logical Depth, the plausible number of computational steps in an object's causal history, is the complexity measure we chiefly recommend. A logically deep object, in other words, is one containing internal evidence of having resulted from a long computation, or from a dynamical process requiring a long time for a computer to simulate. Thus a fossil is deep because it is plausible only as a byproduct of a long evolution, unlike the complementary fracture surfaces in the broken glass example above, which are plausible as the result of a short evolution.

To formalize this notion, we consider the distribution of *running times* of computations by which the standard universal computer might produce the digital output x. Let $P_t(x)$ be the probability that the standard universal computer, randomly programmed by monkeys as before, would produce the output x by a computation that halts in time $\leq t$. Thus $P_t(x)$, for each x, is a monotonically increasing function of t, approaching in the long time limit $P_\infty(x) = P(x)$, i.e. the ordinary time-unbounded algorithmic probability discussed before. A digital object x is said to be "t deep with b bits confidence" iff $P_t(x)/P(x) < 2^{-b}$, in other words, if all but a fraction $< 1/2^b$ of the monkey computations that produce x take more time than t to do so. Inasmuch as the set of universal computations producing x may be regarded as a fairly-weighted microcosm of all causal or logical processes by which x could have arisen, for an object to be t deep with b bits confidence means that the complementary null hypothesis, that x originated by a process of fewer than t steps, can be rejected at the 2^{-b} confidence level, i.e. as less likely tossing b consecutive tails with a fair coin. The confidence parameter b may seem a nuisance, but it is a necessary part of the idea. Since there are many ways of computing any output x, we can make no absolutely certain assertions about how x originated based on intrinsic evidence, only assertions at some level of statistical confidence. As in ordinary statistical discussions, we will sometimes omit mention of the confidence parameter, assuming it to have been set at a value that is safe and conservative in the given context.

Thus defined, depth can be shown to be machine-independent and to obey the slow growth law to within a polynomial in the computation time and an additive constant plus a term of order $\log b$ in the confidence parameter [5]. This imprecision is unfortunately characteristic of the theory of computation times, which typically differ by a small polynomial between one universal machine and another (e.g. one machine may require time $t^2 + 4t + 23$ to simulate what another can do in time t).

Algorithmically random strings, of maximal information content (nearly equal to their length) are shallow because the fast-running print program mentioned above contributes a significant fraction of their rather low algorithmic probability. At the other extreme, trivial nonrandom strings such as '0000000...' are also shallow, because though their algorithmic probability is high, a great deal of it can be accounted for by small fast programs of the form "FOR I=1 TO N; PRINT '0'; NEXT I;". On the other hand a string such as the second million digits of pi, which looks random and is not the output of any known small fast program, but is the output of a small slow program (Compute pi, throw away the first million digits, and print the next million), has the possibility of being deep. (This remains unproven, though. See below for a discussion of provably deep strings.)

Returning to the realm of physical phenomena, we note that use of a universal computer frees the notion of depth from excessive dependence on particular physical processes (e.g., prebiotic chemistry) and allows an object to be called deep only if there is no shortcut path, physical or non-physical, to reconstruct it from a concise description. An object's logical depth may therefore be less than its chronological age. For example, old rocks typically contain physical evidence (e.g., isotope ratios) of the time elapsed since their solidification, but would not be called deep if the aging process could be recapitulated quickly in a computer simulation. Intuitively, this means that the rocks' plausible history, though long in time, was rather uneventful, and therefore does not deserve to be called long in a logical sense.

Although a deep object cannot quickly be made from a shallow one (slow growth rule) a deep object can be quickly made by juxtaposing *two* shallow objects, if these are correlated in a deep way. To see this, let x be a deep string and r be a random string of the same length, generated by coin tossing. Both r and the string y obtained by XORing r and x bit by bit are uniformly distributed over the space of N–bit strings, and so both are with high probability algorithmically random and therefore shallow. However the concatenation string ry, from which x can quickly be made, is deep because of the deep correlation between r and y.

In nature, something like the reverse of this process is more common: a deep object, interacting with its surroundings, typically contaminates them and makes them deep too. For example, outside our hotel, I found this beer-can pull-tab on the ground. I would say that a beer-can pull-tab, although a trivial and worthless byproduct of biological evolution, is so a priori implausible except as a byproduct some such evolution that it probably made the ground it was on nearly as deep as the civilization that produced the beer.

Although time (machine cycles) is the resource closest to the intuitive notion of computational work, space (i.e. memory) is also important because it corresponds to a statistical mechanical system's number of particles or degrees of freedom. The maximum relevant time for a system with N degrees of freedom is of order 2^N, the Poincaré recurrence time; and the deepest state such a system could relax to would be one requiring time 2^N, but only memory N, to compute from a concise description.

Unfortunately, it is not known that any space-bounded physical system or computer can indeed produce objects of such great depth (exponential in N). This uncertainty stems from the famous open P=?PSPACE question in computational complexity theory [11], i.e., from the fact that it is not known whether there exist computable functions requiring exponentially more time to compute than space. In other words, though most complexity theorists suspect otherwise, it is possible that the outcome of every exponentially long computation or physical time evolution in a space-bounded system can be predicted or anticipated by a more efficient algorithm using only polynomial time.

A widely held contrary view among complexity theorists today, considerably stronger than the mere belief that P is not equal to PSPACE, is that there are "cryptographically strong" pseudorandom number generators [6][15], whose successive outputs, on an N–bit seed, satisfy all polynomial time (in N) tests of randomness. The existence of such generators implies that space-bounded universal computers, and therefore any physical systems that mimic such computers, can after all produce exponentially deep outputs.

Deep mathematical objects can be shown to exist without invoking any unproven assumptions by diagonal arguments similar to that used to prove the existence of uncomputable functions. For example, for appropriate values of N (greater than a few thousand, say, to be safely larger than overhead in program size required to combine simple subroutines or program one simple machine to simulate another), the algorithm

```
By exhaustive simulation of all possible computations running
less than $2^N$ steps, find and print out the lexicographically
first $N$-bit string $x$ whose algorithmic probability, from
computations running less than $2^N$ steps, is less than $2^{-N/2}$
```

defines specific N–bit string that by construction is 2^N deep with about $N/2 - \log N - c$ bits confidence, where c is the number of bits required to program the above algorithm in machine language. The string must exist because there are too many N–bit strings for them all to have time-bounded algorithmic probability as great as $2^{N/2}$, and of the ones that do not, there must be a first.

Though such constructions establish the existence of deep objects, actual execution of the algorithm would use so much space and time (exponential and double-exponential in N, respectively) as to be utterly nonphysical.

It is worth noting that neither algorithmic information nor depth is an effectively computable property. This limitation follows from the most basic result of computability theory, the unsolvability of the halting problem, and reflects the fact that although we can prove a string nonrandom (by exhibiting a small program to compute it) we cannot in general prove it random. A string that seems shallow and random might in fact be the output of some very slow running small program, which ultimately halts but whose halting we have no means of predicting. This open-endedness is a necessary feature of the scientific method: at any time some phenomena will always be incompletely understood, so they appear more random and less deep than than they really are.

The uncomputablilty of depth is no hindrance in the present theoretical setting where we assume a known cause (e.g., a physical system's initial conditions and equations of motion) and try to prove theorems about the depth of its typical effects. Here it is usually possible to set an upper bound on the depth of the effect by first showing that the system can be simulated by a universal computer within a time t and then invoking the slow growth rule to argue that such a computation, deterministic or probabilistic, is unlikely to have produced a result much deeper than t. On the other hand, proving lower bounds for depth, e.g., proving that a given deterministic or probabilistic cause certainly or probably leads to a deep effect, though always possible in principle, is more difficult, because it requires showing that no equally simple cause could have produced the same effect more quickly.

Aside from its nonspecific usefulness in clarifying intuition, the notions of complexity discussed here raise potentially decidable questions in statistical physics and the theory of computation concerning necessary and sufficient conditions for the production of complexity, especially logical depth.

In the theory of computation the relation of depth to classic unproved conjectures in time and space complexity has been mentioned.

In statistical physics, the role of dissipation in generating and stabilizing complexity is a major problem area. The need for dissipation to produce and stabilize remote non-additive entropy in locally interacting systems has already been mentioned and is fairly well understood. Concerning depth, one may ask in general how dissipation can help error-correcting computation to proceed despite the locally destructive effects of noise.

One obvious way dissipation assists in error-correction is by allowing compression (many-to-one mapping) of a system's information-bearing degrees of freedom, which, in making the error, have undergone a one-to-many mapping. Another way dissipation may help is by exempting systems from the Gibbs phase rule which applies to equilibrium systems with short-ranged interactions [3]. In typical d–dimensional equilibrium systems of this sort, barring symmetries or accidental degeneracy of parameters such as occurs on a coexistence line, there is a unique thermodynamic phase of lowest free energy. The nucleation and growth of this most stable state renders equilibrium systems ergodic and unable to store information reliably in the presence of "hostile" (i.e. symmetry-breaking) noise. Since they forget their initial conditions, such systems cannot be programmed by them, and so cannot be computationally universal. Analogous dissipative systems, because they have no defined free energy in d dimensions, are exempt from this rule. A $d+1$ dimensional free energy can be defined, but varying the parameters of the d dimensional model does not in general destabilize one phase relative to another [9].

One may ask what other properties besides irreversibility a system needs to take advantage of the exemption from Gibbs phase rule. Known examples, such as Toom's cellular automaton rules [19], lack rotation symmetry, but it is not known whether this is necessary.

Conversely one can ask to what extent equilibrium systems (e.g. quasicrystals) can be caused to have computationally complex ground states, even though they remain subject to the Gibbs phase rule [18].

Finally one can ask whether dissipative processes such as turbulence, that are not explicitly computational or genetic or error-correcting, can still generate large amounts of remote non-additive entropy. Do they generate logical depth? Does a persistent hydrodynamic phenomenon such as Jupiter's Great Red Spot contain internal evidence of a nontrivial dynamical history leading to its present state, or is there no systematic objective difference between a the red spot of today and that of a century ago?

Acknowledgements

These ideas are the outcome of about 25 years of thoughtful discussions and suggestions from from Ray Solomonoff, Gregory Chaitin, Rolf Landauer, Dexter Kozen, Gilles Brassard, Leonid Levin, Peter Gacs, Stephen Wolfram, Geoff Grinstein, Tom Toffoli, Norman Margolus, and many others. Some of this work was done while the author was visiting California Institute of Technology as a Sherman Fairchild Scholar.

Discussion

Schulman Is this "nonadditive entropy" information, entropy, or something you'd measure with a calorimeter?

Bennett Both. It can be expressed as a nonadditivity of algorithmic information, or as a nonadditivity of thermodynamic entropy that could be measured, in the case of the DNA soup, by integrating along a reversible calorimetric path in which the duplicated DNA was reversibly restored to to its non-duplicated state by a carefull reversal of the action of the copying enzymes.

Lebowitz What is the relation between the complexity of the beer can top and that of the Alhambra, or between a Rembrandt painting and a child's crayon drawing?

Bennett The Alhambra is deeper, but maybe not much. Both contain evidence of the general scope of biological and cultural evolution, but the Alhambra may contain evidence of additional causal processes not necessary to produce beer cans, and not likely side-effects of any beer-can-producing civilization.

Miller With regard to your notion of logical depth, what do you mean by "very long"? For what message length does your definition become well-defined?

Bennett The messages need to be longer than the number of bits required to program one simple universal computer to simulate another, or to program the fairly simple algorithms implicit in the proof of the slow growth law, typically a few thousand bits, for depth to be reasonably robust.

Cover You mentioned that algorithmic complexity is computer-independent. Is that true also of logical depth?

Bennett Less so. [As noted above in the printed version of the talk,] logical depth, being based on time complexity, suffers from the polynomial slop typical of time complexity results that attempt to be machine-independent over a reasonably broad range of machines.

Unruh Isn't the function $H(x)$ undefined since you can't know if the random program won't stop.

Bennett This makes $H(x)$ uncomputable, but it is still well defined.

Gell-Mann 1. I believe it would be helpful to include, in the list of systems characterized by what they can do rather than what they are, COMPLEX ADAPTIVE SYSTEMS that can adapt or evolve. 2. Charlie and his friends are typically interested in long messages, for which additive constants and polynomial functions may not matter much. If one cares about systems described by shorter messages, then it is desirable to know from the beginning the describing system, the nature of the language it employs, the coarse-graining of what it is describing, and so forth. Only in that way can absolute quantities be defined, if at all. 3. Although it is not relevant to Charlie's argument, it should be noted that between a schema like DNA and a "phenotypic" object like a human being, a large amount of partly random information is introduced in the course of development, so that the individuality of a human being is much greater than that of the DNA.

Lloyd Is the beer can pull tab as complex as the civilization that produced it?

Bennett It depends on how much of the world's history was plausibly necessary to produce the beer can. It also depends on whether one defines depth using a purely serial machine such as a Turing machine, or a moderately parallel one, capable of simulating, without having to slow down, all the parallel dynamical processes going on in our civilization. In the latter case, the depth of civilization is only greater to the extent that it contains objects not plausible as byproducts of a beer-can-producing civilization, since plausible byproducts could be simulated at no extra cost. In the former case, the difference may be greater, reflecting the extent to which civilization contains evidence of causal processes not plausibly *necessary* to produce beer cans.

Wootters In your definition of algorithmic entropy, is there a reason that you used the monkey formulation rather than the length of the shortest program that produces the desired output sequence?

Bennett Including all the other programs besides the shortest makes only an additive constant difference in algorithmic entropy, but is necessary in the definition of depth, where the other programs, besides the shortest, help to determine the significance parameter. Also it is possible, though not proven, that there may be objects that are "deterministically deep but probabilistically shallow", in other words, objects that have a high fast probability, but no single small fast program.

Zurek First a comment, then a question. Andy Albrecht was wondering about the relevance of such algorithmic considerations to the issue of "coarse grainings," and you have implicated me. I do not want to go into details here, so let me only mention that one way in which algorithmic randomness is helpful in this context is that it can be used to help clarify the well-known problem of the "simplicity" of coarse-grainings. It is often argued that a choice of coarse-graining is a privilege of the observer and, therefore, the entropy is defined with respect to it has an observer-dependent value. This is certainly true. Nevertheless algorithmic randomness could be used to prove that the observers which can communicate with ease will also agree on which coarse-grainings are simple. Therefore their estimates of entropy will agree to very high accuracy. Now for the question. Could you comment on the "thermodynamic depth" which has also been proposed as a measure of complexity?

Bennett I meant to. Thermodynamic depth differs from the complexity measures I have been emphasizing here in that it depends on the history rather than just the state. The thermodynamic depth of the history of the igneous rock would, as I understand, be large, reflecting the large amount of dissipation that occurred in that history, whereas the logical depth of the rock is small, because of the ability to short-circuit this long history by a short computation.

Albrecht What about operating system–dependence? There can be an operating system that prints out the human genome every time you press "H"?

Bennett If the operating system is treated as part of the program (external data fed into the computer) there is no problem. If it is treated as part of the computer, then that computer (with the whole human genome built in) could not fairly be called a simple computer. Even if one perversely decided to use it as the standard universal computer, algorithmic entropies defined on it would not differ from those defined on a simple Turing machine by more than a rather large additive constant, the information content of the human genome.

Albrecht What is the complexity of the system after the bacteria have died?

Bennett Lower. More specifically it depends on how soon after they have died. Immediately afterward, it is probably pretty deep. When the bacteria have all decayed to an equilibrium mixture of carbon dioxide, and water, etc., they are shallow again.

Albrecht So complexity need not increase monotonically like entropy?.

Bennett That is correct.

Teitelboim Is our universe deeper than any other conceivable universe?

Bennett I don't know. I guess that there might be other universes with less wasted motion than ours, more efficient computations and less forgetting of deep things that have been computed before, but on the other hand my remark about deep objects contaminating their environment suggests that not much depth is ever destroyed.

References

[1] C.H. Bennett, "The Thermodynamics of Computation— a Review", *Internat. J. Theoretical Phys.* **21**, 905-940 (1982).

[2] C.H. Bennett, "Information, Dissipation, and the Definition of Organization", in *Emerging Syntheses in Science*, David Pines ed., Addison-Wesley (1987).

[3] C.H. Bennett and G. Grinstein *Phys. Rev. Lett.* **55**, 657-660 (1985).

[4] C.H. Bennett, "On the Nature and Origin of Complexity in Discrete, Homogeneous, Locally-Interacting Systems", *Foundations of Physics* **16**, 585-592 (1986).

[5] Charles H. Bennett "Logical Depth and Physical Complexity" in *The Universal Turing Machine – a Half-Century Survey*, edited by Rolf Herken Oxford University Press 227-257 (1988).

[6] M. Blum and S. Micali "How to Generate Cryptographically Strong Sequences of Pseudo Random Bits," *SIAM J. on Computing* **13**, 850-864 (1984).

[7] G. Chaitin, "A Theory of Program Size Formally Identical to Information Theory", *J. Assoc. Comput. Mach.* **22**, 329-340 (1975).

[8] G. Chaitin, "Algorithmic Information Theory", *IBM J. Res. Develop.* **21**, 350-359, 496, (1977). Cf also *Information, Randomness, and Incompleteness—Papers on Algorithmic Information Theory*, (World Scientific Press, Singapore 1987).

[9] E. Domany and W. Kinzel, *Phys. Rev. Lett.* **53**, 311 (1984).

[10] E. Fredkin and T. Toffoli, *Internat. J. of Theo. Phys.* **21**, 219, 1982.

[11] M. Garey and D. Johnson, *Computers and Intractability, a Guide to NP Completeness* (Freeman, 1979).

[12] P. Gacs, *Journal of Computer and System Sciences* **32**, 15-78 (1986)

[13] P. Gacs and J. Reif, *Proc. 17th ACM Symposium on the Theory of Computing*, 388-395 (1985).

[14] Gardner, Martin, "Mathematical Games" *Scientific American* 20-34 (November 1979)

[15] L. Levin "One-Way Functions and Pseudorandom Generators" *ACM Symposium on Theory of Computing*, (1985).

[16] C. Moore, *Phys. Rev. Lett.* **64**, 2354 (1990).

[17] S. Omohundro, *Physica* (Amsterdam) **10D**, 128-134 (1984).

[18] Charles Radin, *J. Math. Phys.* **26**, 1342 (1985) and *J. Stat. Phys.* **43**, 707 (1986).

[19] A. L. Toom in *Adv. in Probability* **6** (Multicomponent Systems, ed. by R. L. Dobrushin), Dekker, New York (1980), pp. 549-575.

[20] A.K. Zvonkin and L.A. Levin, "The Complexity of Finite Objects and the Development of the Concepts of Information and Randomness by Means of the Theory of Algorithms", *Russ. Math. Surv.* **256** 83-124 (1970).

[21] W.H. Zurek, *Phys. Rev.* **A 40**, 4731-4751 (1989)