Nothing Makes Sense in Computing Except in the Light of Evolution

TOMMASO TOFFOLI

Electrical and Computer Engineering, Boston University

Most work on computation deals with its structural aspects—what it is composed of, how the individual elements work in isolation and how they are connected to one another, how best they can be implemented by a physical systems, etc.—in sum, it concerns itself with a *proximate* account of computation's mechanisms. We argue that, though that kind of work is of course indispensable, computation cannot be understood or even properly *defined* if it is not placed in the context of the evolutionary feedback loop that brought it into being and continually monitors its adaptive fitness with respect to its *ultimate* goals within the organism that employs it.

Far from treating it as an optional, somewhat marginal activity, this viewpoint assigns to unconventional computing a fundamental role in the above evolutionary process, namely, to generate and support through their initial stages those *variations*—those experiments and explorations—that will then be evaluated by evolution's differential-survival engine.

Key words: Computation, evolution, adaptive advantage, dissipation, intentional stance, computation as adaptive trait

INTRODUCTION

Nothing seemed to me more apt, as a banner for the first issue of this *International Journal of Unconventional Computing*, than to paraphrase Dobzhansky's celebrated statement about the unique explanatory power of

---

Corresponding Author: E-mail: tt@bu.edu

Darwinian evolution.[1] A more conventional title for this paper might have been, of course, "Why unconventional computing?" But we have no need to ask rhetorical (and somewhat apologetic) questions. The reasons for unconventional computation come to the forefront by themselves as soon as we answer the really fundamental question, "Why should we (or anything else in the world, for that matter) bother to compute *at all*?"

My plan, then, is to do justice to the latter question first. From the perspective thus acquired we'll then look at representative topics in unconventional computing, noting how much evolutionary history provides a perspective to scientific contents.

THE FOOD OF LIFE

We *eat*, of course, in order to stay alive. (A reminder of this intimate connection between food and survival is given by the English verb "to starve," which, originally meaning just "to die, for any reason," soon took on the more specialized meaning "to die for lack of food.") If we disregard for a moment the role of food as a source of raw building materials—needed only for bodily growth or to replace occasional losses—the chief function of food is to somehow keep the whole works running, day after day. What is food's essential ingredient, and in what manner does this ingredient support life?

In this section I'll argue that food's essential ingredient is *predictability*—a form of entropy. In the next three, that the chief use of predictability is to power life's operative instrument, namely, *computation*.

By the end of the 1700s, industrial machinery with its the ever-hungry furnaces, as well as Lavoisier's quantitative experiments on the metabolism of living things, had made it clear that, like industrial plants, we use food as fuel; that is, as a source of energy. Food, then, is—like money—a *fungible resource*: different forms of it can be interconverted, stored, and traded according to definite exchange rates based on their respective energy contents, and are thus essentially equivalent.

Though its existential consequences were only gradually appreciated, the discovery of the fungibility of food was a veritable "paradigm change," to be compared with the invention of money—the quintessential fungible resource—or even with the invention of *number* as an abstract counting

---

[1]"Nothing makes sense in biology except in the light of evolution." [17].

device.[2] For starters, one could no longer hold the animistic belief that "we are what we eat." If food is just energy, then the Eskimo who lives on an exclusive meat diet for most of the year must be substantially the same human as the Irish peasant who subsists almost entirely on potatoes. (And forget about drawing courage from eating your enemy's heart!) Secondly, outgrowing the Victorian mania for exotica and curios, investigators started looking for what is *common* between forms of food rather than what is different. The moment one starts asking, "We are out of newt's eyes; can I substitute eyes of bat here? And what's all this eyes-of-something stuff, anyhow?" the stage is set for a transition from superstitious medicine to a rational pharmacopea, to a rational materials science and a rational energetics, and eventually to rational theories of information, computation, and even of *soul* [16].

Moreover, an obvious question pops up once one realizes (Lavoisier and all that again) that energy is *conserved*: "But then why are we continually looking for new energy? If energy recirculates, why should one be charged for it?[3]"

In sum, (a) it is not food per se that we are after, as much as the energy that comes with it; and (b) since energy is conserved, clearly it can't be the energy itself that we consume. What we actually seek and consume is a commodity conveyed by certain forms of energy ("high-grade energy," "free energy") and which can be identified with *amount of effective microscopic predictability* (we will justify the qualifiers 'effective' and 'microscopic' in a moment; with this understanding, we'll say just 'predictability' for short).

Predictability turns out to be a quantity of the same kind as what thermodynamicists, after Clausius (1865), had called *entropy*, and, which, according to Boltzmann's bold intuition (1872), has the dimensions of "log of number of states" [9]. Indeed, microscopic predictability is the *same* quantity as physicists call *fine-grained entropy*, but taken with the opposite

---

[2]Most languages still have fossil expressions from a time when different classes of objects had their own grouping methods and counting words. In English, we say "a quire" for 24 sheets and "a week" for seven days; "a brace of partridges," a "pair of shoelaces," and a "yoke of oxen" for different kinds of pairs; and we can order "a gross of pencils" but not "a gross of potatoes."

[3]"…but Edison's greatest achievement came in 1879 when he invented the electric company. Edison's design was a brilliant adaptation of the simple electrical circuit: the electric company sends electricity through a wire to a customer, then immediately gets the electricity back through another wire, then (this is the brilliant part) sends it right back to the customer again.

This means that an electric company can sell a customer the same batch of electricity thousands of times a day and never get caught, since very few customers take the time to examine their electricity closely. In fact, the last year any new electricity was generated was 1937." [Supposed to be by Dave Barry, possibly from The Taming of the Screw; see florence@ipd.info.uni-karlsruhe.de.]"

sign.[4] Besides being an appropriate term for a quantity that points in the same direction as *increasing order*, 'predictability' reminds one that this quantity measures *amount of knowledge* (as to the state of an object system O from the viewpoint of a subject system[5] S), and thus, as any entropy-like quantity, has both objective and subjective aspects.[6]

   Living organisms do need a steady diet of predictability, for a number of related purposes:

1. To *preserve* themselves, that is, to replenish the amount of predictability which is continually leaking out of them. Organisms spontaneously tend to become less and less predictable to themselves, as a consequence both of external disturbances and of the nonlinear propagation of initial internal uncertainties (see the discussion on invertibility below). To compensate for this they take in, as food, materials that have a great amount of microscopic predictability in them (sugar, gasoline, sunlight), and through an appropriate transfer process they perform a *predictability swap* from food to themselves, whereby the food material is left in a less predictable state—and discarded as excrement, garbage, heat—while they are left in a better known state. In other words, organisms perform *error correction* on themselves (not perfect correction, perhaps, but "a stitch in time saves nine!"). Ultimately, many forms of behavior, such as sensing and reacting, defense and offence, and modeling and forecasting, fall in this category. Note that this concept of "active restoration of predictability" is usefully carried over to such natural extensions of an individual's physical body as may be home, car, address book; nest, feeding territory; immediate family, hive sisterhood; and so forth.

2. To take an insurance policy against catastrofic errors by *reproducing*, that is, making multiple backups of their own hard-earned predictable

---

[4]Intuitively, entropy corresponds to 'amount of disorder' (as order increases, entropy decreases). Its opposite, or 'amount of order', has been called negentropy by Brillouin [10] and quite simply information by De Witt and Graham [18]. The term 'negentropy' doesn't seem to have taken root as an English word; the other two, 'order' and 'information', are already much overloaded and may easily lead to confusion.

[5]The latter, of course, need not be a human; an insect, a steam engine, or a cryptographic device are legitimate subjects.

[6]"Our probabilities and the entropies based on them are indeed 'subjective' in the sense that they represent human information; if they did not, they could not serve their purpose. But they are completely 'objective' in the sense that they are determined by the information specified, independent of anyone's personality, opinions, or hopes. It is 'objectivity' in this sense that we need if information is ever to be a basis for new theoretical developments in science." [24, p. 390].

selves onto appropriate "writable media." This is done by taking in raw materials having a plausible elemental composition and rearranging their atoms[7] so as to assemble an organism structurally and functionally similar to the original one. The criterion for 'sufficient similarity' is whether the new organism will be able to reproduce in turn—and so forth recursively.

3. To *fuel reliable computation*. Even in the presence of an overall favorable entropy gradient the elementary processes involved in 1 and 2 would hardly ever spontaneously take place at an appreciable rate. To move along they have to be "chaperoned" step-by-step by appropriate control machinery that sequences and coordinates the desired reactions while inhibiting other equally plausible but undesired ones. It is this kind of intermediate activity of a general-purpose "managerial" nature, as contrasted to the "production" functions—physical manufacturing and maintenance—of points 1 and 2, that goes under the name of *computation*.

    Like all forms of management, computation in any of its forms incurs *substantial operating costs*. In spite of its abstract nature, computation's costs are *of the same kind* as those accrued by concrete functions such as pumping water uphill. In fact, an organism's manufacturing and management departments are both financed by one and the same source—the *predictability* treasury.

    Owing to its attention to "whether" and "when" and "how many times," to "which" rather than "what," to identity or complementarity between two objects (like a key and a lock) rather than the specific nature of either ("is the relevant pattern a shape, a smell, or a gesture?"), computation is intrinsically an abstract information-processing activity. One generic "technological kit" of symbolic tokens[8] and logic operations can be used to manage an extraordinary variety of complex physical functions.

    The rest of this paper is one long argument aimed at showing that the role of "management in behalf of life" we have just outlined is computation's very *definitional essence*.

---

[7] Or, sometimes, even larger ready-to-use building blocks such as certain simple molecules.

[8] The essence of a symbolic system is that the initial assignment of logic roles (like 'true' and 'false') to physical tokens (like spin-up and spin-down) is in principle arbitrary — a "broken symmetry" — but once made becomes, through an uninterrupeted historical chain, a shared inheritance of all users — what's 'true' for one is 'true' for all.

INVERTIBILITY AND THE SECOND PRINCIPLE OF THERMODYNAMICS

Before proceeding with our main discussion it will be useful to anticipate two questions which address subtler aspects of the predictability economy.

The first is, "Why do we have to make recourse to this order-for-disorder swap? (See point 1 in Section 2.) Can't errors be simply *erased* right where they are?''

No, they can't! Suppose that a certain system variable q has three possible states, 0, 1, and 2, and that in the present circumstances it is ought to be in state 0. If we are not sure of q's actual state we can try to correct a potential error at this stage by applying the rule "If q is in state 0, leave it alone; otherwise, change its state to 0," as per the following diagram:

$$0 \atop 1 \longrightarrow 0 \atop 2 \qquad . \tag{1}$$

However, it turns out that the microscopic dynamics of our physical world is, to all evidence, strictly *invertible*:[9] distinct states always flow into distinct states—they never *merge* as in (1). Intuitively, in an invertible system a full memory of the past is preserved in any future state.[10] Such a world is literally "informationlossless;'' in it, fine-grained entropy (and likewise microscopic predictability) are strictly conserved. That means that a diagram like the above can't be the whole picture; somewhere in the system there must be some other variable, say, p, such that the three-way merge for q as in (1) is accompanied by a three-way *split* for p. Therefore, any increase in certainty in the value of q, as achieved for instance by (1), must be paid for by a proportionate increase in the uncertainty of the variable p.

Remembering that it was the purported loss of predictability (in one's body and immediate environment) that created the need for food in the first place, as an immediate riposte to the above answer[11] we get the

---

[9]That is, it's a one-to-one correspondence between current state and next state. This holds whether we stick to the Lagrangian dynamics of classical physics or the unitary dynamics of quantum physics.

[10]Much as in a deterministic system the future is fully determined by the past, so is in an invertible system also the past fully determined by the future.

[11]When we'd said, in essence, that predictability can only be imported, not created.
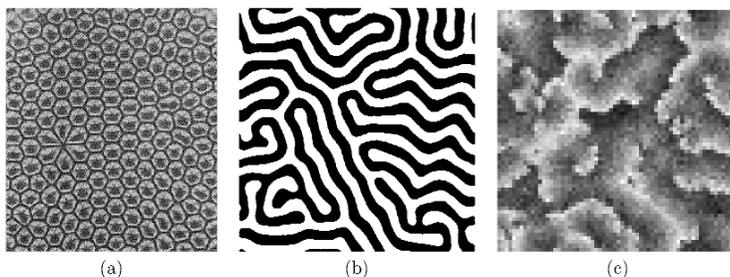
FIGURE 1

Dissipative cascades: (a) Bénard cells. These convection cells arise and organize themselves spontaneously in a shallow liquid uniformly heated from below and losing heat from the upper surface fast enough for the temperature there to be lower than at the bottom. (b) Magnetic domains arising from the contrast between ferromagnetic effects, which would tend to align all the spins in the same direction and lead to a single ('up') or ('down') domain, and classical magnetic forces that would tend to have adjacent magnets counter-aligned in a head-to-tail fashion. (c) "Forest fire," or traveling self-sustaining waves arising from the weak coupling of an array of identical relaxation oscillators.

second question, "But, if predictability can't be lost *at all*, why should one need to replenish it?"

This paradox is resolved if we recall that what we seek to restore is not how much can be predicted by a hypothetical Laplacean demon[12] having some initial information about the *entire* universe and all the computational power and the time in the world, but how much can be *effectively* predicted by a living organism which, with limited resources and in an adversary environment, must make timely life-or-death choices. (A bicycle secured by a 5-digit combination lock is in principle unprotected: any amateur Laplace can "predict" the right combination by a trial-and-error process of moderate length and walk away with the bike. But in practice the bike is reasonably safe: a real burglar will find it more rewarding to go for an easier pick, and in any event won't want to become too obvious a pick himself.)

In a later section we shall have more to say about how a competitive regime affects an organism's efficiency in husbanding its predictability budget.
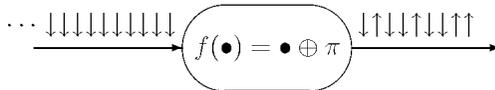
---

[12]"An intellect which at any given moment knew all the forces that animate matter and the mutual position of the beings that compose it, if this intellect were vast enough to submit that data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom: for such an intellect nothing would be uncertain; and the future just like the past would be present before its eyes." [Pierre Simon, Marquis de Laplace (1749-1827), Philosophical Essay on Probabilities (1814), SpringerVerlag 1995]

IS THIS A COMPUTATION?

Coming back to the main discussion, we may observe that computation seems to be everywhere—but what is it precisely? A common way to tell a computation is by what Aristotle would have called its 'efficient cause'—what today we'd call its *proximate mechanisms*. From this viewpoint, the distinguishing feature of a computation is a network of *signals* and *events*. Though this network may be of large or indefinite extent, yet both signals and events are drawn from a limited repertoire, so that in the long run events of the same kind and are bound to appear over and over, and similarly for signals. Think of a mechanical clock, crammed with springs and gear; the schematics, at the gate-and-wire level, of an electronic calculator; or biochemical machinery and products such as those of Fig. 2. In sum, think of any mechanism that can produce a *boundless variety of effects* out of a *substantial uniformity of means*. (The prime exemplar of such a family of mechanisms is, of course, the *Turing machine*.)

For instance, the following processes would likely be recognized as computations:

- The inner works of an ordinary general-purpose computer.
- The emergence of complex structures from within a *dissipative cascade*,[13] as illustrated by Fig. 1.
- A cell phone's managing to "find its place" in the network, recognizing and guiding the progress of a call, and providing clerical services (menus, address book, etc.) to the user.
- A complex mechanical coupling between a low-entropy input and a still low-entropy but logically deeper[7] output, such as that achieved by a device (depicted below) which, when fed with a stream of spins all pointing down, flips some of them so that the sequence of ups and downs in the output stream corresponds to the binary digits of $\pi = 11.0010010\ldots$ (cf. [8]).[14]

$$\cdots \downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow\downarrow \quad \overbrace{f(\bullet) = \bullet \oplus \pi} \quad \downarrow\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow\uparrow\uparrow$$

---

[13]I.e., a medium which, sandwiched between an energy source and a thermal sink, is made active by the entropy flow from one to the other.

[14]Note that, in principle, the latter could be an invertible device, not requiring a power supply to operate, and could equally turn the (perfectly predictable) sequence of digits of p into a more practical "fuel" such as a string of all 0s.
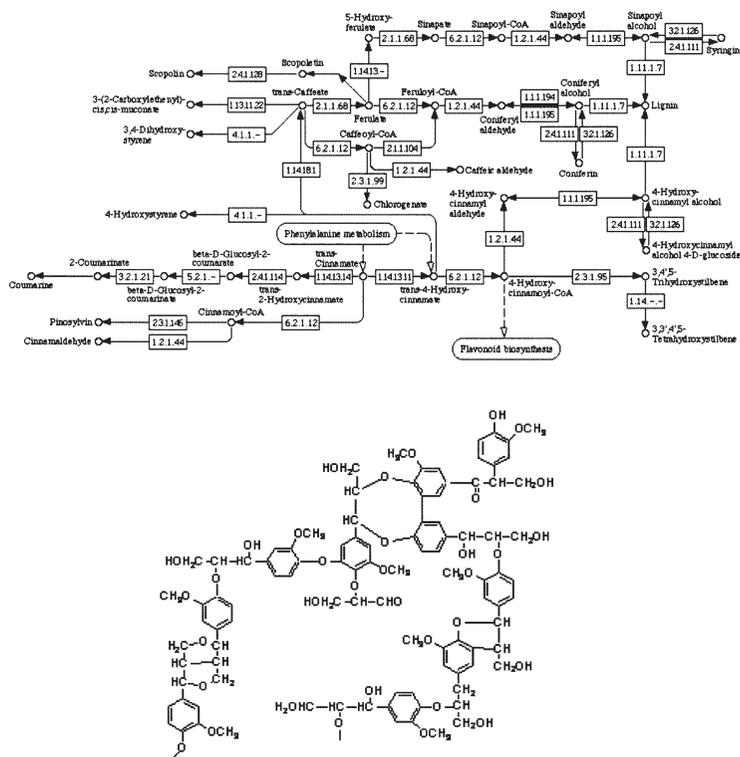
FIGURE 2
At the top, chemical pathways of stilbene, coumarine, and lignin biosynthesis [27]. At the bottom, a small section of the resulting lignin polymer illustrating some typical chemical linkages created by the reactions on the left [27]. Note how, much like in an Erector set, both the process and the product use repeated instances of the same few building blocks and interconnection patterns.

- The conversion, by plant cells, of atmospheric carbon, oxygen, hydrogen, and sunlight into fibers of lignin (Fig. 2).

In all the above examples we have the two prerequisites of (a) uniformity of building blocks and (b) open-ended variety of behavior. I do not deny that, given the appropriate context, any of the above examples should count as a 'computation'. What I claim is that these are necessary but not sufficient conditions. A third, essential, requisite besides the substance and form of the physical process is (c) a specific kind of *context*, which we'll introduce at the end of the present section and elaborate in the following one.
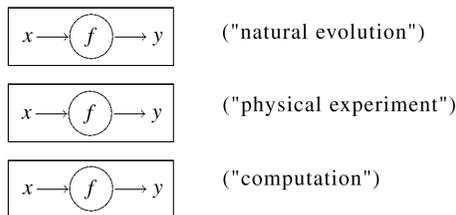
To let you see for yourself that something more than bare operating

FIGURE 3
(a) "Tête de garçon III" (Pablo Picasso, 1962). (b) "Face on Mars" (natural rock formation in the Cydonia region, Mars, 1976).

mechanism is needed to characterize computation, I shall invite you to point at an ongoing physical process—a waterfall, a soccer game, a portion of a living cell, the inside of a carburetor, a rock on the beach—and tell me, case by case, "This is a computation," "This is not a computation," or, if you need to be more nuanced, "This aspect of this phenomenon may be viewed as a computation." How can you tell? How would you explain to somebody else how to tell? For that matter, on what grounds would you try to explain to the browser of checkout-counter magazines that the "tête de garçon" on the left (Fig. 3) is a face while the "face on Mars" on the right is not?

To put things in an even starker light, I'll challenge you to tell me the difference between the three boxes below.

$$x \longrightarrow \boxed{f} \longrightarrow y \qquad (\text{"natural evolution"})$$

$$x \longrightarrow \boxed{f} \longrightarrow y \qquad (\text{"physical experiment"})$$

$$x \longrightarrow \boxed{f} \longrightarrow y \qquad (\text{"computation"})$$

In each of the boxes, a system of initial state x undergoes a specified dynamical transformation f and re-emerges in a final state y. In the first box the system follows its "natural evolution"—and no one need to care or even know about x, f, and y ("If a tree falls in the forest and there is nobody around does it make a sound?"). In the second box, what we are uncertain about is the dynamics f; to collect information about it we perform a "physical experiment," i.e., we prepare the system in a state x, we let it evolve through f, and then examine the final state y; from the correlation

between x and y we try to infer the nature of f. After a variety of such experiments we may become so confident in our knowledge of f that we are willing to use the apparatus to perform a "computation" (third box), that is, to predict or construct y by means of f from a given x.

But the boxes themselves are identical! It is only their intentional context that makes us label them differently.

To summarize, it is true that in ordinary circumstances most computing machinery—mechanical, electronic, or biological—can quite confidently be recognized by its structural features. In fact, a conventional definition that I've often used myself runs as follows:

> "Computation is the exercise of *function composition* in a context where the building blocks—*the data storage*, *transport*, and *interaction* primitives—(a) are *finite objects* and (b) exist in a *finite number of types* specified once and for all."

More intuitively, computation is a discipline in which to achieve novel functions one is not allowed to introduce *novel types* of components, but only to appropriately combine together *more instances* of components drawn from the same, fixed and finite, repertoire. Note that this is just an extension from the data storage to the data processing realm of what Schrödinger—having DNA in mind—termed an "aperiodic crystal" [37].

The problem with the above definition, based only on efficient causes, is that it encompasses just about *any kind* of activity that displays a certain loose regularity or repetitiveness—snowflakes, systems of waves or vortices, polycrystalline materials, cellular automata, etc. In fact, it includes anything that can take place in a world governed by *uniform laws*—a world, that is, where the same kinds of primitive components are found and the same composition rules apply at any place and any time. Even partial differential equations—the very kind of rules that until recently were considered the format *par excellence* for the laws of physics—fall under the scope of that definition. It is true that PDEs use uncountable infinities both in their coordinates and in their state variables, and thus strictly speaking violate the local finiteness requirements (a) and (b) above. However, since they are locally linearizable, they can be simulated to any desired degree of approximation not only by algorithms operating on a discrete mesh, but, using a sufficiently fine mesh, by algorithms such as cellular automata that use *finite-state machinery* at each site [42].

Shall one then indiscriminately call computation any composite physical phenomenon that satisfies a certain finitary discipline, namely, a system whose components and composition rules are drawn from a bounded

repertoire? Though I'm sympathetic with the idea of viewing the entire physical world, governed as it is by uniform laws, as "one big computer," this form of reductionism is not very useful for our present purpose. In the 1700s, the planetary system was likened to "a cosmic clockwork;" that is, the workings of the universe were interpreted in terms of the most complex mechanism that at that time humans could *make, control, and wholly comprehend* themselves. What we want to do here is not help explain physics by analogy with the familiar computer—today's "clockwork"—but try to pin down the essence of computation itself. And that cannot be done by concentrating on structure alone, leaving out those aspects that have to do with overall function and goals (and which Aristotle would have called 'final causes'). In brief, we'll have to bring into the picture the intentional stance [15].

SYMMETRY BREAKING AND EVOLUTION: MORE ME!

We've just seen that it is not useful to call 'computation' just any nontrivial yet somewhat disciplined coupling between state variables. We also want this coupling to have been *intentionally* set up for the purpose of predicting or manipulating—in other words, for *knowing* or *doing* something. This is what shall distinguish bona-fide computation from other intriguing function-composition phenomena such as weather patterns or stock-exchange cycles. But now we have new questions, namely, 'Set up by whom or what?", "What is it good for?', and "How do we recognize intention?"

Far from me to want to sneak animistic, spiritualistic, or even simply anthropic considerations into the makeup of computation! The concept of computation must emerge as a natural, well-characterized, objective construct, recognizable by and useful to humans, Martians, and robots[15] alike.

Fortunately, much of the necessary conceptual apparatus has been available in a tentative form since Darwin, and has recently been so well developed and consolidated (I'm talking about the past twenty-five years) that it is now capable of addressing, as *ordinary scientific questions*, some issues that philosophers had long been struggling with but were no longer expecting, after numerous inconclusive attempts, to see resolved soon. I'm referring to such topics as *causality* [30] and *inference* [24], *consciousness*

---

[15]Here I'm especially thinking of Jaynes's robot [24], a device proposed, like Maxwell's demon for thermodynamics and the Turing machine for calculation, to help get to the core of an issue (in Jaynes's case, inference) by divesting it of many human assumptions and unknowns.

and *free will* [16,31], *intention and design* [15,11-13,36], the *spontaneous emergence of complexity* [25], *the nature of knowledge* [32], and *altruism*, the "*good of the group*," and even—why not?—*love* [48]. What's surprising is that in virtually all these cases the breakthrough was made possible by positing Darwinian evolution—the differential survival arising from error-prone replication—as an underlying mechanism, and then working out case-by-case its long-term implications.

Differential survival[16] can easily be conceded because it sounds innocuous. Far from being implausible it is, in an appropriate setting, almost tautological. Since it does not have to be endowed with foresight it does not threaten religious sensibilities or our unicity as intelligent beings. And, in any event, experimental evidence for it, often no harder to obtain than by "kitchen biology," is now pervasive. Its long-term implications, on the other hand, may catch one unprepared: Darwinian evolution seems able to create from nothing—or at least to mimic in an uncannily deceptive fashion—intention and design.

We should be ready by now to take the plunge. Instead of assuming that intention, consciousness, and all the other goodies mentioned above, including—and this is my only original contribution to the present argument—computation, have a pre-defined existence (or a pre-existing definition?) in some vague and as yet *unidentified* world of Platonic ideas, and grudgingly admitting that evolution is somehow doing a rather good job at *counterfeiting* them, why don't we proceed in a more natural direction and eliminate both mysteries (the two italics right above) at a stroke? Let us agree that cause, intention, knowledge, etc.—these resilient and useful bundles of behavioral features—are *by definition* to be identified with certain common and well-characterized outcomes of any protracted evolutionary process, and that the animistic "cause," "intention," "knowledge," used in common human parlance are but informal shorthand for the corresponding evolutionary constructs!

"If we accept this," you may argue, "instead of learning about Darwinian evolution in a biology textbook we'll soon find that life is a chapter in an evolution textbook!" Certainly! And you wouldn't be the first to arrive at that conclusion:

> "If you took a vote today, the most popular definition of life
> would probably be the one proposed 10 years ago by Gerald
> Joyce of the Scripps Institute in La Jolla, California. He describes

---

[16]I'm deliberately using a more neutral term than, say, the badly loaded 'selection of the fittest'.

life as a self-sustaining chemical system capable of evolving
through Darwinian natural selection." [Bob Holmes, "What is
life," *New Scientist* **182**:2463 (4 Sep 2004), 32]

Before saying more about this let me give an example, right from
computation's backyard, of an analogous—and successfully completed—
reversal of direction. *Entropy* was introduced by physicists to facilitate
energy accounting in analyzing the operation of steam engines. In 1848,
even though it was not clear precisely what kind of stuff it was, entropy was
in any event viewed as a quintessential physical quantity like energy and
mass: you could keep it in a box, you could measure it in good *physical*
units of joule/kelvin, you could siphon it from one container to another, you
could couple it to other physical quantities by plugging it into certain
differential equations. We'll let Boltzmann, Gibbs, Einstein, Bose, Dirac,
and many others work on the set for a century, gradually changing the scene
from Thermodynamics to Statistical Mechanics. When we raise the curtain
in 1948, entropy is still in physics' court. Shannon has been working on a
quantity having to do with the probability distribution of messages and
which happens to appear in a mathematical formula like that used by
Boltzmann for entropy [38]. "Would it be confusing,'' he asks von
Neumann, "to call this quantity 'entropy' because of its analogy with
physics?" Allegedly von Neumann gives him the go-ahead: "First of all this
quantity of yours does indeed remind one of entropy, and in second place no
one knows what entropy is[17], so in a debate you will always have the
advantage.'' Shannon of course follows his advice.

Now, Shannon's entropy is a perfectly well defined and well-understood
numerical quantity associated with a particular kind of *lack of knowledge*
called a 'probability distribution.' It is an abstract mathematical quantity
that arises out of sorting and counting—mountain heights, text syllables,

---

[17]Note the analogy with traditionally defined consciousness, for one, which everybody uses
rather consistently even without really knowing what it is [15].

[18]A old stumper was the Gibbs paradox — that one may arrive at different answers for the
entropy of the same piece of material by following different measurement strategies — is easily
dismissed by the "new" entropy [23]. Since entropy is, as we've seen, a quantity associated
with a lack of knowledge, the entropy I compute for a lump of material depends on exactly what
I let the computation know about what I know. In most cases, the entropy of a lump of matter
depends almost exclusively on the parameters of its standard thermodynamic description and
only marginally on different ways of preparing the sample—much as the amount of tax due in a
typical IRS return depends much more on my salary, about which I cannot lie since it is known
to the IRS, than on little things I may want to risk getting away with. Nonetheless, one can
always concoct cases where not "reporting" certain items may leads to different results, and
there is no paradox in that.

licence plates, runs of heads and tails from a biased coin . . . never mind what—*anything* that can be sorted and counted, and this includes energy states of atoms. And when it is used for atoms, its value happens to coincide with that of physical entropy.

At this point you will not need my encouragement to propose that one may well kill two birds with one stone and declare that the mother of all entropies is Shannon's entropy, henceforth to be called simply 'entropy', while physical entropy is just a special case of it, suited to its more restricted context. If physicists happened to stumble on it first it's because their subject matter forced them to deal, well before biologists or accountants, with the problem of sorting and counting truly astronomical number of states. We are certainly grateful to physicists for their pioneering work, but entropy—thank you very much!—truly belongs to *everyone*.

Moreover, since after this reversal entropy is clearly a human invention, not a material ore that physicists mine, we (humans) know perfectly well what goes into it and may claim that we *do* now know what entropy is. Physical entropy may still offer some 'intelligence problems[18]" or computational challenges, but its nature is no longer a mystery.

Coming back to the proposal to interpret will, love, and "all that" as "derivatives" (as in 'financial derivatives') of evolution, what remains to be done is

1. Show that Darwinian evolution can be characterized as a purely mathematical construct, much as "continuity," "eigenvalue," "entropy," or "feedback loop," and from which certain types of derived constructs can be obtained in a systematic way, much as from the concept of entropy one can derive that of maximum-entropy distribution, and, from that, systematic prescriptions for optimal inference, for locating equilibrium points, etc.

2. For each specific behavioral trait or phenomenological feature to be explained as a consequence of evolution, such as "hypocrisy" or "group selection," show the *mechanism* by which this feature would effectively come into being or be "synthesized" by the very fact of being adaptively advantageous under a Darwinian evolution regime. No mechanism, no admission to the club!

Point 1 is by-and-large an ongoing collective enterprise that has already made much progress [25] and is continually being refined (e.g. [39]). I will explain the general trend in the rest of this section. The responsibility for point 2 falls, of course, on the entrant, as admirably explained by Wilson [48]. I do some of that pleading in behalf of *computation* at several places in this paper.

FIGURE 4
Arctic fox: summer and winter fur. [Photos: www.bbc.co.uk/nature/wildfacts]

Evolution is in effect a general-purpose servomechanism for solving problems which are hard to solve but for which it is easy to determine whether an alleged solution is indeed a solution. (The NP-complete problems that play such an important role in the theory of computation complexity belong to this class.) How does such a mechanism work?

We are all familiar with the duality between *structure* and *function*. Philosophers and biologists introduce another useful distinction—that between *proximate* and *ultimate* explanations. For example, the arctic fox molts twice a year, in the spring to a brown fur and in the fall to a white one (Fig. 4).

The *proximate* explanation is that hormones released by certain patterns of daylight length stimulate a complex chain of chemical processes like "inhibition of a specific protein A releases the production of protein B which catalyzes a certain reaction C" and so forth. A hypothetical breeder, who may want to induce an earlier fall molt so as to get a richer white fur, will care about all the details of that explanation. The fox, or, better, "foxhood," doesn't. What really counts is the *mimetic* function: a brown fox is easier to spot against a snowy background and is thus more likely to miss a meal or, even worse, become a meal. This is an *ultimate* explanation; if we are mathematically oriented we may think of it as the *equivalence class* of all those proximate mechanisms or structures that yield a certain functional trait. Since foxes that didn't bear this trait didn't leave as many descendants,

---

[19]Darwin, matter-of-fact as usual, speaks of "descent with modification" rather than "evolution.".

this trait is *advantageous*; since those who did got it from parents that had a similar trait and will leave a similar trait to their offspring, the trait is transmissible and modifiable[19], or *adaptive*.

Foxes are easy to visualize. But, in general, what kind of entity can be the beneficiary of an adaptively advantageous trait? Can it be an organism? An individual? A species? An intelligence? The answer—which I'm afraid is going to be anticlimactic—can only be "a data structure that, as the outcome of a *broken symmetry*, happens to be what it is but might easily have been otherwise;" in other words, *the memory of a macroscopic historical accident that is continually regenerated by dissipative means*.

Precisely because of the underlying symmetry—the one that was broken—most historical accidents are unenduring. The floatsam pattern left on the beach by one wave is dissolved and recomposed into a new pattern by the next wave; we can think of the train of waves as a "pump" feeding energy into the dissipative process that constitutes the floatsam patterns' world. A special pattern, however, may happen to enter into some sort of resonance with the energy pump, and establish a positive feedback loop that tends to give the pattern a higher degree of *permanence*. It is in this sense that we speak of "advantageous" manipulation of the environment by part of the pattern: the advantage—or the "success," if you like is,—*by definition*, the permanence itself.

Note that we seek permanence of the data pattern, not necessarily of the physical substrate. This permanence may be sought by means of *pre-emptive* defense against attacks to the pattern, or through after-the-accident *correction of errors*, or, finally, through *self-replication*—the hope in this case being that in the long term at least one good copy of the pattern will remain extant.

Even though all three approaches are in fact used, it is the last one that seems to afford the best strategy for permanence—it is more general, more flexible, and, even as it entails more drudgery and waste, it requires less "brains," as it were. But it comes with a surprising fine-print contract; in fact, as we shall see in a moment, it is a *Faustian bargain*.

Symmetry breaking with runaway amplication. On a majority-vote election, a certain district will be conquered by the Blue party. Slightly different details in the background noise, like the spread of certain rumor, followed by runaway amplification, might let the district fall into the hands of the Red party instead. But once the district is Blue, much of the administrative machinery will come to be controlled by the Blues and will

be servoed to the consolidation and spread of Blueness. The 'me' in "More me!'' is, in this case, Blueness itself.

Descent with modification—but randomized, please. An amoeba reproduces by gently splitting into two *identical* amoebas—no color distinguishes the two. In spite of this true (rather than broken) symmetry, from the viewpoint of one of the two daughters—initially distinguishable from the other merely by its geometrical coordinates—there is a world of difference between herself and her sister. One is "Me, here!'' and the other is "Non-me, there!'' The two sisters may well be identical; yet, even if in this sense they are of an "equal'' mind, they are not, by construction, of *one* mind. And, as we well know from human history, mere geometrical otherness is one of the most common initiators of symmetry breaking. When resources are limited, others like me—*precisely* because they are like me and desire the same things—become my worst enemies. In this predicament, an amoeba may reason[20] as follows. "Let's face it—I'm not smarter or stronger than the other guy; in fact, as far as I know the situation is materially symmetrical and I only have *less* than a fifty-fifty chance of getting out of here alive (less, since of course there is a third outcome of nonnegligible probability, namely, the death of *both*). What if I changed my structure just a little bit, so as to break the symmetry myself instead of relying on nature's dice?[21] I'm not likely to decrease my chances, and I may even increase them a little.'' By symmetry, one should expect that the other amoeba will come up with the *same* strategy; and this would essentially bring them both back to the starting point unless—as game theory teaches— the "same'' strategy they both use is a *randomized* one and thus may yield, on any individual trial, *different* outcomes for the two parties. In that case, if the randomness generator is good, the small changes that the two amoebas make on themselves will be *independent*. Now the two amoebas will end up being *different*, after all, and may avoid mutual destruction (if one easily prevails on the other we have a *constant-sum* game and the chances attain fifty-fifty) or may even have developed a taste for *different resources* and peacefully go their own different ways.

The Faustian bargain is that, precisely in order to maximize its own chances of survival, an amoeba accepts that what will come out of this

---

[20]Here of course I'm not referring to an individual amoeba's proximate "thoughts" but using a customary metaphor for the ultimate outcome of the evolutionary process; in brief, I'm taking refuge behind the intentional stance [15].

[21]Remark that this approach is not unlike settling out of court to avoid a judge's uncontrollable way of "breaking the symmetry.''

process is no longer (an identical copy of) itself but something a *little different*. To paraphrase the Sermon of the Mount (Matt. 16:25), "Whoever would save his life, will lose it; and whoever loses his life…, he will save it." A related aspect of this Faustian bargain is that, because of its inherent symmetry, intraspecific competition makes it pointless in the long run to push for more muscle, higher speed, better patronage, or more intelligence than your neighbor, since saturation of resources will soon be reached in this constant-sum game [45]. On the other hand, more computing power used to diversify or "discorrelate" yourself from your neighbor may end up being slightly advantageous to *both*:[22] the game no longer need be constant-sum. To add a line to the Ecclesiastes (Eccl. 9:11), "And I saw under the sun that the race is not to the swift, nor the battle to the strong, neither yet bread to the wise, nor yet riches to men of understanding, nor yet favor to men of skill; but time and chance happen to us all" "*And precisely to break this uniformity will I turn to computing power for my strength and support*!"

All the same, any weakness in one's randomness generator can in principle be exploited by the adversary, as players of Morra or "rock paper scissors" well know; this will put a selective pressure towards harder-to-guess randomness generators. (If all this sounds too theoretical or contrived, look at the substantially similar problems in the damn serious business of sexual recombination [35], where the symmetry is almost perfect and the stakes high, or at the multilevel randomizing strategies used by the immune system.) Here we already see a glimpse of those "adaptive forces" that will tend to reward organisms capable of deploying more powerful computing machinery.

EVOLUTION IS ANYTHING BUT PARSIMONIOUS

In the Holmes quote near the beginning of the previous section, life was defined as a self-sustaining dissipative process of a *chemical* nature subject to Darwinian evolution. Holmes was obviously thinking of ordinary earthly biology. By removing the condition 'of a chemical nature' we would extend the definition to include life based on other physical domains such as fire patterns or hydraulics, or on more abstract substrates such as an ordinary digital computer (e.g. Tom Ray's Tierran ecology [34]) or a cellular automaton (e.g. the self-replicator sketched by von Neumann [46]). By then imposing

---

[22]Think of the randomizing algorithms that are used to minimize Ethernet collisions on successive retries without requiring coordination or cooperation.

certain regularity conditions as explained in section "Is this a computation?" we would home in on computation, which is thus characterized as a specific appurtenance of life. Namely, as anticipated at the end of the second section ("The food of life"), computation may be thought of as the *management department* of a living entity (or of some subsystem of it, such as the immune system of vertebrates). Finally, in the same section we said that computation uses up a good fraction of the predictability budget.

Note that physicists and computer scientists have come to the conclusion that the mechanics of computation can in principle be carried out in a quasi-reversible, virtually dissipationless fashion [5,6,19,41]. That might seem to imply that computation doesn't need after all to be fueled in any substantial way. Unfortunately, to achieve a good approximation of this ideal computing regime one must deploy a much larger number of components—gates and wires—than the minimum necessary. There is a close analogy here with *recycling*: to consume fewer raw resources one can recycle most materials—paper, steel, etc.—but at the cost of introducing more temporary-storage bins, reprocessing plants, transportation trucks, management and information-processing capabilities, etc. Moreover, the amount of extra infrastructure needed for this grows faster and faster as one aims to recycle a greater fractions of waste materials. Finally, besides requiring a one-off capital investment, this infrastructure also incurs day-by-day operating expenses proportional to its extent. A point is soon reached where more thorough recycling becomes anti-economical.

Furthermore, in the case of computation, to spend less fuel on a given task one must *decrease the operating speed*. Soon the fitness cost of having to wait a longer time for the results will offset the fuel savings. Adaptive fitness tradeoffs will determine how "quick and dirty" an organism will have to be. In many circumstances, even obstentatious waste can be fairly adaptive.

I'll finish by examining some examples of unconventional computing approaches in the light of the above evolutionary considerations.

BIRTH AND DEATH OF A THINKING MACHINE

A brain is an expensive perk: ours, with 2% of body weight, hogs 20% of the oxygen supply (and this is only one type of computation going on in our body). Other species make do with much less brain—evolution's backroom

accountant keeps reminding each organism that, from their respective niche in the world, the benefits of more brains would not be worth the costs.

Clearly, our species has discovered new uses, for the brain, that bring outstanding benefits. Concurrently, it has found a way to reduce its biological costs. Our evolutionary fast lane is in technological development rather than DNA upgrading, and this includes *artificial*—that is, technology-based—intelligence, or AI. "The ultimate effort is to make computer programs that can solve problems and achieve goals in the world as well as humans." (This and the next few quotes are from the AI pioneer John McCarthy [28]; emphasis mine.)

In 1957, AI and economics pioneer Herbert Simon[23] predicted that within ten years a digital computer would be the world's chess champion [4]. In reality, it took *forty* years for this prediction to come true—and I doubt that Simon could truly anticipate how much more power computers would have gained by then.

How much computing power do we need for artificial intelligence, and of what kind? And how do we find out?

"After WWII, a number of people independently started to work on intelligent machines. The English mathematician Alan Turing may have been the first. He gave a lecture on it in 1947. *He also may have been the first to decide that AI was best researched by programming computers rather than by building machines.* By the late 1950s, there were many researchers on AI, and most of them were basing their work on programming computers."

"Many researchers invented non-computer machines [unconventional computers], hoping that they would be intelligent in different ways than the [conventional-]computer programs could be. *However, they usually simulate their invented machines on a* [conventional] *computer and come to doubt that the new machine is worth building.* Because many billions of dollars have been spent in making computers faster and faster, another kind of machine would have to be very fast to perform better than a program on a computer simulating the machine."

---

[23]In 1975, Simon earned the Turing Award for his work in computer science; in 1978, the Nobel Prize in Economics.

[24]"On May 11 1997, DEEP BLUE defeated Garry Kasparov in a 6-game match held in New York. This was the first time a computer defeated a reigning world champion in a classical chess match." [47].

[25]Advanced Research Projects Agency, established 1958. Then DARPA (1972), then back to ARPA (1993), and then (1996) back to DARPA again.

Obvious as Turing and McCarthy's wisdom (in the emphasised passages above) may be, it would be naive to expect evolution to follow only "wise" paths, that is, paths that are most efficient from an overall, long-term viewpoint (and whose view, for that matter?). In the sixties and seventies, claim making and computer development in AI had stabilized on rather predictable pattern. "Yes, compared with our three-year proposal," the PI would say (a typical interlocutor would be the daring, patient, and deep-pocketed Uncle Sam agency ARPA[25]), "today's demonstration is only a toy example, but what can you expect with only 4 Kilobytes [*sic*] of core memory? Let us have 64 Kbytes, and then you'll see . . ." A year later, the ante would be upped to 256 Kbytes, then one Mbyte, and so forth. The story would then become, "If we only had a compiler[26] we would be able to write and modify much larger programs!" So was born FORTRAN. "If we only had a computer language suitable for AI tasks!" And LISP was born, and other programming languages and tools. "Tapes are so slow—we really need a hard disk." "Ten Megabytes is not enough; we'll soon need one-hundred Mbytes—no, wait, make it one Gigabyte, . . . ten, . . . one hundred—and then we will show you what AI is really capable of!"

Finally, with much ongoing progress in neural networks (stimulated and to a large extent delivered by—who would you expect?—theoretical physicists [20]), relational databases, and associative memories, the appearance of experimental multiprocessors and parallel processors, and the nascent theory of complex systems[27], the AI community started clamoring for a medium capable of massively-parallel fine-grained computation. "We can't really do much in the way of intelligent behavior until we have computing stuff that's organized more or less like the brain—plenty of tiny processors, a lot of generic, redundant interconnections, and no need for detailed point-to-point wiring schematics." And what about programming? "Oh, the stuff will program itself, out of extensive autonomous interaction with its world, out of massive systematic searches through the space of all conceivable actions, or perhaps with the help of deliberate training by means of representative stimulus/response sequences."

So went the Connectionistic Gospel [22,29]. And, since no commercial outfit was about to deliver this magic hardware, somebody from the prime AI laboratory in the country set about designing it

---

[26]A program to turn a high-level computer language into executable machine language.

[27]How complex behavior can emerge from simple parts, provided one is allowed to use as many as desired.

himself, attracting ample public and private funding. So was born the Connection Machine [21]. Conceived, built, programmed, and marketed by the cream of AI researchers, you could have expected that, as soon as this machine was out, every AI lab in the world would line up—waving, of course, million-dollar bills in their hands—and scramble to be the first to grab a unit. No such thing happened—you can read the rest of the story in [40]. Please do.

May Danny Hillis have made a horrible blunder and ended up building the *wrong* machine? I'm not so sure of that. Something like the Connection Machine was what everyone wanted—or at least *said* they wanted—and in this sense, with the wisdom of the moment, it was the *right* machine for somebody to build. What Danny perhaps didn't realize is that he was calling a bluff. It is one thing to say, "We have this beautiful research plan that we'd like to get funding for; we'll only be able to work on the theoretical aspects, of course, until the right hardware becomes available.'' It is quite another thing to be suddenly confronted with the hardware *we'd been clamoring for*, and be asked to demonstrate on it—*right now*—our theory's worth.

At this point, somebody may ask how expensive calling a bluff, wittingly or otherwise, should be allowed to be, and who should foot the bill—but that is another story. It has to do with honesty, justice, social responsibility, all things that deeply affect the human individual but evolution can be quite cavalier about. This issue is admirably addressed by Richard Dawkins [14]. *Exactly* because we live in a world governed by nature's indifference, we have the duty to counter that by using our gift of foresight and try to avoid so much pain, waste, and blunder. Evolution won't—and by its very nature *can't*—do that.

On the other hand, in terms of the affairs that evolution does care about, the Thinking Machines episode was ordinary administration. How many dates doesn't a palm have to produce to insure a reasonable chance of breaking even—of still having one descendant in a few generation? How many projects does DARPA have to finance to insure that once in a while something as successful as the Internet will take off? How would you go about deciding (in 1984) whether connectionism is (a) God's gift to the world, (b) an idea whose time hasn't come yet, (c) a bad idea, (d) a technology that will be invaluable but only for a few niche applications? and so forth. Recalling the characterization of evolution given just before Fig. 4 (solutions that are hard to find but easy to verify), evolution's banner is "How can you find out if you don't try?''

LATTICE-GAS HYDRODYNAMICS

Besides, many other ideas were evaluated in concomitance with the Connection Machine.

I remember the feeling of exhilaration I had when I saw the stately sloshing of a lattice-gas fluid on my cellular automata machine screen [43]. No one has seen that before! I had taken a cellular automaton scheme that Norman Margolus has set up for studying microscopic computational processes and I had "perversely" programmed it with a rule whose behavior would only make sense from a macroscopic viewpoint. I knew next to nothing of hydrodynamics, and so was not the best person to send into the world this creature I had stumbled on. Yet it was with some regret that, in doing a bibliographical search, I discovered that something similar had already been proposed a couple of years before—as a purely theoretical conceptual joke from a professional hydrodynamicist (Yves Pomeau [33]) to his professional hydrodynamicist colleagues, and then of course immediately let drop.

A few months later we invited Pomeau to a miniconference and showed him these lattice-gas fluids. According to Pomeau himself, seeing those simulations running live on our cellular automata machines made him realize that what had been conceived primarily as a *conceptual* model could indeed be turned, by using suitable hardware, into a *computationally accessible* model: this stimulated his interest in finding lattice-gas rules which would provide better models of fluids. A landmark was reached with the slightly more complicated FHP model (it uses six rather than four particle directions) which gives, in an appropriate macroscopic limit, a fluid obeying the well-known *Navier-Stokes* equation, and thus suitable for modeling actual hydrodynamics. Soon after, analogous results for three-dimensional models were obtained by a number of researchers.

By that time, Stephen Wolfram, who also had been impressed by these simulations and had been drawn into the area of cellular automata himself, had started applying his outstanding mathematical physics expertise to lattice gases. Finally, the Connection Machine had already failed to make a splash first as an AI tool and then as an executive's perc, and was now looking for new markets. As a design afterthought, the Connection Machine had been provided with a number of fast numerical-processor "islands" dispersed through a sea of fine-grained processors. How about turning it into a scientific computer? And so was born a collaboration between Hillis, Wolfram, and others to turn the Connection Machine into a lattice-gas

hydrodynamics computer. The notion was not without merit, and actually produced a number of experiments that were impressive for that time, but eventually proved not to be competitive enough.

The reason? Cellular automata carry within them the seeds of their own destruction, as it were. By painstakingly simulating the gyrations of millions or billions of tokens, they are ideal for deriving reliable mesososcopic predictions directly from somewhat stylized microscopic principles. But this effort literally exhausts their computational resources, so that none are left to go one step further and synthesize fully macroscopic behavior. However, the mesoscopic parameters and properties obtained once and for all by means of cellular automata models can in turn be plugged into different kinds of models, which bridge the gap between mesoscopics and macroscopics. Like Moses, who died on Mount Nebo in sight of the Promised Land after having guided the Chosen People for forty years trough the desert all the way from Egypt, and had to let a new generation take the final step, so cellular automata's best exploits take place in the wasteland between micro- and mesoscopics, while the information so gained goes on to be used by others.

As a matter of fact, a shadow of the former lattice gases is retained to advantage in the move to the new, meso-to-macro, models, The latter, which use so-called *lattice Boltzmann* algorithms, skim the best from several approaches and yield industrial-strength results. A viable corporation, EXA (www.exa.com), has after all sprung out of all this trouble.

Aren't the ways of evolution hard to fathom?

CONCLUSIONS

Computation is that component of life that pursues control tasks through regular, symbolic mechanisms; it is only recognizable within the "pseudo-purposive" feedback loop set up by evolution. In this context, unconventional computing is a natural expression of the principle that, as in war and love, "all is fair in evolution."

REFERENCES

[1] Adamatzky, A. (ed.), *Collision-Based Computing*, Springer 2002.

[2] American Academy of Arts and Sciences, *Artificial Intelligence*, special issue of *Proc. Am. Acad. Arts and Sciences* **117**:1 (Winter 1988).

[3] American Academy of Arts and Sciences, *A New Era in Computation*, special issue of *Proc. Am. Acad. Arts and Sciences* **121**:1 (Winter 1992).

[4] Andresen, S,, "Herbert A. Simon: AI Pioneer," *IEEE Intelligent Systems*, July/August 2001, 71-72.

[5] Barenco, A., Bennett, C., Cleve, R., DiVincenzo, D., Margolus, N., Shor, P., Sleator, T., Smolin, J., and Weinfurter, H., Elementary gates for quantum computation. *Phys. Rev*. A **52** (1995), 3457-3467, 1995

[6] Bennett, C. H., Logical reversibility of computation, *IBM J. Res. and Develop*. **6** (1973), 525-532.

[7] Bennett, C. H., Logical depth and physical complexity, T*he Universal Turing Machine—A half-century survey* (Herken, R., ed.), Oxford U. Press 1988, 227-257.

[8] Bennett, C. H., The thermodynamics of computation—A review, *Int. J. Theoret. Phys*. **21** (1982), 905-940.

[9] Boltzmann, L., Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. den Säatzen über des Wärmegleichgewicht," *Sitzungsber. Kais. Akad. Wiss. Wien Math. Naturwiss. Classe* **76** (1877), 373-435.

[10] Brillouin, L., *Science and Information Theory*, 2nd edition, Academic Press 1962.

[11] Dawkins, R., *The Selfish Gene*, Oxford U. Press 1976; 2nd edition 1990.

[12] Dawkins, R., *The Extended Phenotype*, Oxford U. Press 1982.

[13] Dawkins, R., *The Blind Watchmaker: Why the evidence of evolution reveals a universe without design*, Norton 1987.

[14] Dawkins, R., *A Devil's Chaplain*, Houghton Mifflin 2003.

[15] Dennett, D., *The Intentional Stance*, MIT Press 1987.

[16] Dennett, D., *Consciousness Explained*, Little, Brown and Co. 1991.

[17] Dobzhansky, T., Nothing makes sense in biology except in the light of evolution. *American Biology Teacher* **35** (1973), 125-129.

[18] Everett, H., The theory of the universal wave function, in *The Many-Worlds Interpretation of Quantum Mechanics* (De Witt, B. and Graham, N., ed.), Princeton U. Press 1973, 3-137.

[19] Fredkin, E., and Toffoli, T., Conservative Logic, *Int. J. Theor. Phys*. **21** (1982), 219-253; reprinted in [1], 47-81.

[20] Hertz, J., Krogh, A., and Palmer, R., *Introduction to the Theory of Neural Computation*, Addison-Wesley 1991.

[21] Hillis, D., *The Connection Machine*, MIT Press 1985.

[22] Hillis, D., *What is massively parallel computing, and why is it so important?* in [3], 1-15.

[23] Jaynes, E., The Gibbs paradox, in *Maximum-Entropy and Bayesian Methods* (Erickson, G. et al. ed.), Kluwer 1992.

[24] Jaynes, E.T., *Probability Theory—The Logic of Science*, Cambridge University Press 2003.

[25] Kauffman, S., *At Home in the Universe: The search for laws of self-organization and complexity*, Oxford U. Press 1995.

[26] Kyoto University Bioinformatics Center, Stilbene, coumarine, and lignin biosynthesis, *Kyoto Encyclopedia of Genes and Genomes*, www.genome.jp/kegg/-pathway/map/map00940.html 2004.

[27] Lignine Institute, Lignine and its properties, www.lignin.info/01augdialogue.html.

[28] McCarthy, J., What is artificial intelligence? http://www-formal.stanford.edu/jmc/whatisai, 29 March 2003.

[29] Papert, S., One AI or many?, in [2], 1-14.

[30] Pearl, J., *Causality: Models, Reasoning, and Inference* (reprinted with corrections edition), Cambridge U. Press 2000.

[31] Peres, A., Existence of 'free will' as a problem of physics, *Found. Phys.* **16** (1986), 573-584.

[32] Plotkin, H., *Darwin Machines and the Nature of Knowledge*, Harvard U. Press 1996.

[33] Hardy, J., de Pazzis, O., and Pomeau, Y., Molecular dynamics of a classical lattice gas: Transport properties and time correlation functions, *Phys. Rev.* **13A** (1976), 1949-1960.

[34] Ray, T., How I created life in a virtual universe, www.his.atr.jp/~ray/pubs/nathist 1993.

[35] Ridley, M., *Mendel's Demon: Gene justice and the complexity of life*, Orion 2000. Republished as *The Cooperative Gene: How Mendel's Demon explains the evolution of complex beings*, Free Press 2001.

[36] Schrage, M., Revolutionary evolutionist, *Wired* 07 (July 1995), www.wired.com/wired/archive/3.07/dawkins.html.

[37] Schrödinger, E. What is life? (1944), in *What Is Life? and Other Scientific Essays*, Doubleday 1956, 1-88.

[38] Shannon, C., A mathematical theory of communication, *Bell Sys. Tech. J.* **27** (1948), 379-423 and 623-656.

[39] Stewart, I., How the species became, *New Scientist* **180**:2416 (11 Oct 2003), 33.

[40] Taubes, G., The rise and fall of Thinking Machines, *Inc. Magazine*, September 15, 1995.

[41] Toffoli, T. , Bicontinuous extensions of invertible combinatorial functions, *Math. Systems Theory* **14** (1981), 13-23.

[42] Toffoli, T., Cellular automata as an alternative to (rather than an approximation of) differential equations in modeling physics, *Physica D* **10** (1984), 117-127.

[43] Toffoli, T., and Margolus, N., *Cellular Automata Machines: A new environment for modeling*, MIT Press 1987.

[44] Toffoli, T., and Margolus, N., Invertible cellular automata: a review, *Physica D* **45** (1990), 229-253.

[45] Toffoli, T., What you always wanted to know about genetic algorithms but were afraid to hear, *Festschrifts in honor of John Holland*, (Booker, L., et al. ed.), Center for the Studies of Complex Systems, The University of Michigan (June 1999), 131-136.

[46] von Neumann, J., *The Theory of Self-reproducing Automata* (Burks, A., ed.), Univ. of Illinois Press 1966.

[47] Wall, B., "Computer chess history" www.geocities.com/SiliconValley/Lab/7378/comphis.htm, 18 July 2004.

[48] Wilson, D., *Darwin's Cathedrals: Evolution, Religion, and the nature of society*, The U. of Chicago Press 2002.