What Automata Can Provide a Medium for Life?

Tommaso Toffoli(✉)

Boston University, Boston, MA O2215, USA
tt@bu.edu

Abstract. Hadn't this question already been answered? We all know about computation-universal Turing Machines. And we know that any such machine can simulate a space-time dynamics not unlike von Neumann's cellular automaton, which is computation- and construction-universal and among other things can play host to self-replicating machines. And that self-replication sprinkled with a bit of randomness should inexorably lead to descent with variation, competition, and thence to evolution and all that.

And note that the state of the art has much advanced in the fifty years since. "So?" Enrico Fermi would have asked, "Where are they?"

It turns out that life is by its very nature a marginal, fragile, and ephemeral kind of phenomenon. For a substrate or a "culture medium" to be able to support it, computation- and construction-universality are necessary—but by no means sufficient! Most automata (including, I suspect, Conway's very game of Life) will go through their entire life course without ever originating anything like life.

What questions, then, should we ask of a prospective medium—be it a Turing machine, a cellular automaton, or some other kind of automaton—that will probe its capabilities to originate and/or sustain some form of life?

1 Introduction

In this paper I dwell on certain concerns that in my opinion haven't been voiced loudly enough—if and when they were articulated at all. These concerns have to do with the *spontaneous emergence* and the *continued sustenance* of something like *life*. More specifically, if *lifelike* behavior is desired to emerge out of an *automaton-like* medium—for definiteness, a distributed dynamical system such as a Turing machine or a cellular automaton—what conditions may make this emergence and its historical persistence more (or less) likely? To wit, *what kind of dynamical laws* and *what kind of initial conditions*?

A first-year student of theoretical computer science will of course be tempted to answer my question by means of a simple, though grandiose, reductionistic plan. "What's the problem?" she may ask. And explain:

M. Cook and T. Neary (Eds.): AUTOMATA 2016, LNCS 9664, pp. 10–25, 2016.
DOI: 10.1007/978-3-319-39300-1_2

1. We have an uninterrupted four-billion-year historical record of the emergence of Life-as-we-know-it here on Earth, an average planet of an average sun of an average galaxy. The latter is one of the few dozen galaxies that make up the Local Group galaxy cluster, which in turn is an average component of the Laniakea Supercluster. The observable universe comprises ten million odd such superclusters (en.wikipedia.org/wiki/Supercluster). Thus Earth is an undistinguished place; you're likely to find one like it in the vicinity of no matter where you look in the universe.
2. We know physics—the dynamics of our universe—at a level of detail sufficient to account for the properties of ordinary matter, run-of-the-mill cosmology, the laws of geology and chemistry, and biological structures and processes.
3. We have a reasonable idea of what the state of our universe could have been like, say, ten billion years ago, well before our solar system (and thus Sun, Earth, earthly life, and all that) had emerged.
4. Finally, I just learned that we can construct a *Universal Turing Machine* (UTM). To make a long story short, according to the Church–Turing thesis, any ordinary computer, including the minuscule Intel 4004—the first microprocessor, with only 2000 transistors (today we have microprocessors with 4 billion transistors!) are *Universal Computers*. That is, they can be programmed so that, given enough memory and time, they can exactly emulate (though perhaps more slowly and less efficiently) anything that can be computed by any other physical computer.

"Given all of that," our student will continue, "I would just program my computer to run a *simulation model* of our universe. Points 2 and 3 above will give me all the laws and data I need to write my program. Even if from my simulation I don't get exactly the kind of life that we have here on Earth today, I should get something in that ballpark, since by Point 1 an environment like Earth is so common. At worst, to find life I would have to look at several places and/or try several times! Thus the procedures and parameters of my computer program, which embody Points 2 and 3 of our physical world, do specify *for an automaton* dynamical laws and initial conditions of a kind that make the emergence of life in it likely, as per your request."

If we accept the Church–Turing thesis (which I do), my student's answer will be formally correct, but only insofar as tautological—and thus essentially vacuous. In effect she would be saying, "If our universe produced life [which I'll grant], and if my simulation by means of an automaton-like medium is functionally equivalent to the running of that physical universe itself as per Points 2 and 3 [which I'll grant as well], I don't see how this automaton would fail to often come up with life."

"Fine," I would say, "If we ignore for a moment that yours is just a thought experiment—your simulation would need astronomical resources to be carried out, including maybe trillions of "universe times"—then what you are saying is correct. *Unless* . . . " "Unless what?" "Unless the emergence of life is so fiendishly dependent on the laws of physics being *just so* and the initial conditions being *just so*, that any small deviation from that would risk making us miss the sweet spot—Turing and Church notwithstanding!"

In sum, to discover (or learn) what kind of dynamical laws and initial conditions are favorable to life we need an approach that is not only more *practical* but also more *robust*. We may eventually have to accept that life depends in an *essential way* on the fine-tuning of a certain parameter, but it would be perverse and wasteful—and, I'd say, non-scientific—to start from the outset by looking for an explanation of life that depended on the fine-tuning of multiple parameters. We'd rather study different properties that address different features of life, or produce some aspects of it in isolation. Once we've gained familiarity with one property (which may be of an ON-or-OFF kind or come with adjustable parameters) and with the effect of the constraints it imposes on a system (since studying a property *is* imposing a constraint), we can try to combine two or more of those properties and see how much interesting space is left as their intersection, if any. Echoing one of Richard Feynman's last messages, "What we cannot create we don't understand."

I had the fortune to live through and personally take part in a discovery process (concerning the computation and construction capabilities of *invertible cellular automata*), that followed by and large the above strategy; this is chronicled with copious references in [5]. At the end of the day, Norman Margolus and I could conclude, "Only a few years ago, what was known about invertible cellular automata could be summarized in a few lines—and wasn't very exciting either. Today, one can tell a [much] more interesting story."

What is remarkable is that there we'd asked a question of automata theory—"Can general computation take place in an *invertible* automata medium?"—at a time when there were conjectures and (alleged) proofs that the answer would be "No!" But we said, "Wait a moment! Physics (whether classical or quantum) seems to be able to do computation under the same constraint (of invertibility) and the additional constraint of *continuity*. What is going on? How can physics do that? How can we emulate it in automata? And what (if any) do we have to give up in return?" In the end, the answer turned out to be "Yes"—we'd been able to 'throw out the bath water' without losing the 'baby'!

2 Some Historical Notes

The challenge of creating contrivances able to imitate life and thought has been recognized since antiquity, as the myths of Daedalos, Pygmalion, and the Golem attest. The increasing technological sophistication of the 1700s and 1800s, with Vaucanson's dancing and flute-playing automata (1737), Jacquard's programmable tapestry loom (1801), Charles Babbage's differential engine (1830) and his more ambitious analytical engine (the first general-purpose computer, left incomplete, 1850), Louis Braille's writing machine (a true dot-matrix printer, 1839), and innumerable other creations, both responded to and fueled a popular fascination for ever more capable lifelike mechanical creations. At the same time, it seemed evident that higher levels of behavioral complexity could only be achieved by introducing ever more sophisticated, more precise, and fundamentally more *capable* or *powerful* mechanical parts. For instance, wind-up springs were successfully replaced, as a source of motive power, by hydraulic and pneumatic actuators, by steam engines, and eventually by electric motors.

That more complicated behavior is only made possible by more capable components has been a commonly accepted view for much of human history. This perception was philosophically formalized by Aristotle (ca. 330 BC) and (in a Christian-adapted version) by Thomas Aquinas (ca. 1270). The latter argues that stones are ... just stones—they don't do anything and don't need much of an explanation. But plants grow, reproduce, and survive. For this they obviously need a special faculty—let us call it a *vegetative* soul. Animals also grow, survive, and reproduce, but in addition they move on purpose, see, feel, react, and communicate. For this, they obviously need a higher-level faculty, which one may call a *sensitive* soul. Humans do all of the above, and also have abstract thought, articulate speech, and long-term planning. For this—you guessed it—they must be endowed with an even higher-level faculty, which one may call *rational* soul. One thus postulates a hierarchy of faculties as an "explanation" of an observed hierarchy of behavioral complexities. Neat, but vacuous. I'm sure that if one wanted to explain the faculties of angels one could always ascribe them to an ad hoc *angelic* soul (incidentally, Thomas Aquinas' moniker in the trade was "Doctor Angelicus").

Sporadic attempts were made in the past to replace this ever-ascending hierarchical ladder of "faculties" or "souls" with a single, unified, reductionistic once-and-for-all hypothesis. We shall just mention Democritus' atomic theory (400 BC), later picked up by Epicurus (ca. 290 BC) and masterly elaborated by Lucretius, in his *De Rerum Natura* ("On the make-up of things," ca. 90 BC), into a comprehensive and humane naturalistic doctrine. But, as we've seen, the Aristotelian approach, which described things in a way closer to what they superficially like from a human standpoint—and thus mentally less demanding—remained until recently the more widely circulated. Under the guise of Neo-Thomism, it was given official sanction as a philosophical framework for catholic theology (by Pope Leo XIII's encyclical *Aeterni Patris*) as late as 1879!

One could quibble whether the promotion, say, from "vegetative" to "sensitive" was a fundamentally *discrete* step, requiring the infusion of a higher-level soul, or could be achieved by a gradual, *continuous* process, but it seemed evident that the gap between "vegetative" and "inert," between living and non-living, was a major one, unbridgeable without external intervention. Indeed the distinction between "inorganic" and "organic" chemical compounds originally rested on whether they could be synthesized in the laboratory or only by living organisms. This gap was famously breached in 1828 by Friedrich Wöhler's discovery that urea—a byproduct of animal metabolism—can be produced in the lab from inorganic starting materials, thus helping put to rest the widely held doctrine of "vitalism."

As is well known, an analogous breakthrough occurred in the 30's, with independent but convergent results by Gödel, Post, Church, and Turing. Computation can be mechanized, and moreover reduced to primitive mechanical components of utter simplicity—basically, just wires and NAND gates. Granted that faculty, and thus anywhere above a very low threshold of *quality*, as it were, computability is just a matter of *quantity*—how long your program has to be to describe how to compute a certain function in a given programming language. You don't have to add new constructs to your language—or new modules to your microprocessor—to compute a more complex function, or, for that matter, *any* computable function. The *intelligence*—the "white collar"—needed for a computation can be distilled and captured once and for all onto memory as a text program. The repetitive, low-skilled *labor*—the "blue collar"—is provided by a microprocessor built once and for all and which will do for *any* task. To have the computation performed, one just has to provide as much of those raw, unstructured, "passive" resources—energy (or food), memory space (or papyrus), and time—as required.

Think of the swarm of human calculators organized into a "hive computer" by Richard Feynman during the Manhattan Project. A numerical computational task would be broken down into subtasks that individuals of modest mathematical skills could carry out with pencil and paper—some sort of glorified accounting (have you ever seen a 'calculus pad'?). To perform their job, these human calculators didn't even have to know what the overall product would turn out to be. This was actually a *design feature* of the whole outfit—you cannot accidentally let out a secret if you were not told more than you strictly need to know!

Just as Turing & Co. had "mechanized thought" in the 30's, in the late 40's John von Neumann tackled the problem of "mechanizing life." He had many talents as a mathematician, physicist, and systems scientist. The precarious alternation of success and tragedy for Jews in Central Europe must have imprinted him with a frantic urgency not only to seek personal security as a "court magician" (his productivity as an indispensable scientific consultant for government and enterprise was immense), but also to investigate in a systems-theoretic way the problems of self-preservation and self-reproduction.

So from models of abstract thought he turned for a brief time to models of material life—movement, action, construction, growth. self-reproduction. Steering him away from models based on differential equations and moving mechanical parts, his colleague Stan Ulam convinced him of the practicality of using instead automata-like models of space, time, and dynamics. So were born *cellular automata*—a new incarnation of discrete recurrence relations—within which he sketched a strategy for designing self-replicating spacetime structures. As soon as such a strategy appeared viable (see Sect. 3.1), he turned to different projects. His orphaned project remained in the care of his assistant (the scholar Arthur Burks) who eventually completed it and wrote an extensive reasoned report about it [9].

It is a symptom of von Neumann's urgency that, in spite of his being a towering figure in theoretical physics, his cellular automaton substrate is used for an empirical—macroscopic and phenomenological—model. He tried to achieve his goal, self-replicating structures, with the "greatest economy"—that is, the "cheapest and dirtiest"—of means. He put very little theoretical physics in his cellular automata—just local interactions in a space and time framework. No energy, no inertia, no action and reaction, no invertibility, no thermodynamics or second law. In this sense, cellular automata were not only left as orphans, but also deprived of much of the rich inheritance they could have expected from such a father. It took another forty to seventy years for this neglect to be remedied [3–5].

I'm not trying to be a curmudgeon. Von Neumann *created from scratch* something that shows *some* of the essential aspects of life (we'll have to say more about this). Our problem, however, is different ("What laws and initial conditions are likely to lead to the *spontaneous emergence* of life within an automaton?") and presumably much harder. Its solution would be proportionately more rewarding. It might even provide useful suggestions for the preservation as well as the expansion of our Earth's life experiment.

3 Methodological Problems

Here I should discuss a number of methodological problems, illustrating each by one or two examples.

3.1 von Neumann's Fixed-Point Trick

I will start by giving credit to von Neumann on a point where most credit is due to him.

An important aspect of von Neumann's strategy is that his creatures happen to be *self-replicating* even though they are not literally *self-copying*. It is to his credit that, as a systems scientist, he quickly homed in onto one of the most robust and dependable strategies for self-reproduction, anticipating the discovery of ribosomal function by Zamenicnik [12] and the structure of DNA [11] by Watson and Crick [11]. The famous last sentence of [11], "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material," was a noncommittal way of stating a conjecture that involved the entire cell-reproduction scheme.

In a nutshell, to reproduce you should not try to copy your *phenotype* (your body); that is hard, impractical, and unreliable. You rely instead on your phenotype to already contain a *genotype* (a recipe or set of instructions) for building *some* phenotype (not necessarily like yours) and equipping it with a copy of the original recipe itself as a genotype. Besides containing this recipe, your genotype should be able to read it and carry it out, so as to construct a new complete phenotype. If this "daughter" phenotype itself happens to be able to read and execute its recipe, then its own daughter will happen to be identical to a copy of itself—even though this daughter was not *copied* but *made from a recipe*—and so forth, generation after generation.

With this process, you may not have ended up reproducing yourself (though after the second generation your descendants *will* reproduce *themselves*). However, if the genotype carried by phenotype X *happened* to be a recipe for the very phenotype X that carried it, then self-reproduction will start with the first generation. In the latter case, the phenotype cum genotype happens to be a *fixed point* of a *recurrence relation* (the chain of generations).

Our problem with von Neumann's brilliant insight is merely that it is not very relevant to the *first stage* of the emergence of life—though it becomes extremely relevant at later stages. (In fact, one may wonder whether *any* later stages of evolution would gamble with doing without it.)

To come to the point, observe that von Neumann's self-replication strategy requires a *matched combination* of an active-fabricator phenotype and a passive-recipe genotype. For the strategy to be successful, that genotype must *happen* to be just what is needed to direct the fabrication of that very phenotype. This requirement is clearly a tall order, as it requires a lot of pre-ordained coordination, and thus is not likely to have been the reproduction mechanism used right at the origin of Earth's life (or of any other emergent life). This mechanism must have been derived later, by a succession of cautious "make before break" steps (cf. [2]).

3.2 Computation Universality Is Not Enough

Computation universality of a medium is essential for it to support life, but that does not mean that just proving the medium's universality by embedding a Universal Turing Machine (UTM) into it makes it particularly likely to support the spontaneous emergence of life in such a way that I can observe and study it.

The argument, already sketched in the Introduction, is simple. It is true that if you put a UTM in my computing medium and show me how its tape is realized within this medium, and similarly for the machine's head, and explain to me how the machine's states are encoded in the head and the tape symbols on the tape, and finally show me the head's transition table (its "microcode"), then I'll be able to program this machine. With an appropriate programming and data-entering effort, I should be able to instruct this machine to simulate, say, a 7-dimensional cellular automaton, a neural network, a chess-player self-learning machine, or, to any required approximation, a continuous mechanism like the solar system. In principle I could even program the UTM to simulate *our entire universe* to study the emergence of life in it, perhaps with a slow-down factor of only 10^{30} and on a scale of 1,000,000,000:1 (so that I'll have to borrow another universe with a volume 10^{27} times as large as mine just to provide the machine's tape, and I would have to hibernate myself and wake up every few billion years just to see one second's worth of simulation progress. But after only a few of those seconds my universe, the UTM in it, and myself will have already vanished in a puff of smoke!

Of course if my computing medium is a high-definition 3-D cellular automaton that can run a Giga-step/s without requiring the power of a huge hydro-electric plant, and if I can easily manipulate its state and visualize its contents, then I might be able to run on it certain focused experiment that I could not easily perform on a conventional computer. But, at that point, I would be *very* surprised if its designers hadn't thought of designing computation universality into its mechanisms at the outset—at virtually no extra cost.

Computation universality is so cheap that it would be a crime not to have it everywhere, but its real worth is as a conceptual tool—it doesn't magically turn a circuit of a dozen gates traveling up and down a tape into an hypercomputer.

3.3 The Fragility of "All Mine!"

Let's consider a luxuriant millenary oak forest. It is conservatively estimated that the average oak will have produced in its lifespan 10 million acorns. How many of these (averaging over all trees) will grow into a new oak? [Think by yourself for a moment; the answer is given a couple of paragraphs down].

Lisp in its pure form is an autistic's dream. It neither produces nor feels any side-effects. A pure Lisp program is a static hierarchical tree of text and its execution leads to the growing in real-time, by iteration and recursion, of a dynamic execution tree of procedure calls and data. The Lisp program is a "Master of the World." It can branch out and out indefinitely, reaching out and spreading at will into an infinite empty space that is ready to be occupied and colonized. There are no competitors, no partners, no superiors. No noise, no disturbances, no unforeseen events, nothing beyond one's control. No alien footprints on the beach!

The essence of life, on the other hand, is competition—not in the sense of the Nietzschean superman, but in the humbleness of presuming approximate *symmetry* with our competitors (the "Copernican principle"), in the recognition that for everyone to be above average is a contradiction in terms. In a world of all heroes, no one is a hero! In life, success in competition mostly entails being adept at sharing, collaborating, and compromising. The reason for all this is very simple—we live in a polynomial spacetime, not an exponential one! In our world, the only long-term exponential growth is that with base one: $1^n \equiv 1$.

To come back to the oak-and-acorns puzzle. This is a millenary forest, so a climax forest—one that has attained dynamic equilibrium. For any acorn to grow into an oak, one oak must have died (lightning, disease, old age) to leave room for it. The lavish investment of 10,000,000:1 is to insure that "the moment one parking spot gets freed, one of your proxies will be right there to take it before somebody else does."

The above considerations are aptly illustrated by the three examples below. Scenarios that were purportedly designed to be "life-ready"—and for all I know they still might be—but within which "life" was mostly sought by means of a Nietzchean approach. The latter, as we hinted above, was bound to founder.

von Neumann's self-replicators. In von Neumann's cellular automaton, an initial self-replicator will construct a second one next to it; after this, the first will go dormant while the second will build a third one, and so on, eventually resulting in a whole row of frozen structures except for an active "tip." What will happen if we plant in this landscape a second "seed" at some distance, oriented perpendicularly to the first? When the tip of this new row impinges on the first row, a collision will occur whose consequences will depend on the relative space and time phase of the two rows and on the actual rules that are in force (von Neumann never completed his project, and many of his followers devised different variant completions [10]). There are of course in von Neumann's project no provisions for his self-replicating automata to deal with obstacles, to fight with one another, to repair themselves, to reuse scattered materials. Depending on a balance between the construction and destruction powers given to different parts of the machinery, we may imagine that the typical resolution will be a clogged up jam, a structure that shrinks until it disappears, or a runaway structure that spreads uncontrollably.

Of course one can resign oneself to using a single row of these self-replicating structures and adapt it to behave like a Turing Machine, in which case one has total control over it *as a Turing Machine*, and we know one can simulate a whole universe with it, just as we saw in the Introduction, and yet, for the same reasons discussed there, be nowhere nearer to getting an insight into the emergence of life—in spite of the self-replicating capabilities of the elements that make up the substrate of this Turing machine. But what is the point, then? The realizations described in [10] mention *billions* of time-steps for a single replication cycle! Whatever computing I need to do, I'd rather use a general-purpose computer by default, and turn to a dedicated processor if I indeed need to simulate a special architecture like a neural network or a cellular automaton.

To effectively use a von Neumann-like cellular automaton as a tool for research on the emergence of life, one will first have to plan (or let evolve) a whole consistent *ecology* that can take advantage of that architecture (even if only at the level of mental experiment). Otherwise all one will achieve is the typical outcome of impatient kids playing with a "chemistry kit." Stinks, explosions, broken glass, the more interesting chemical run out first, and eventually mom will come, say "Why don't you go and play outside?" and proceed to clean up the mess.

Conway's Game of Life. Much more attention to developing an empirical experimental playground was given in Conway's game of Life. It has simpler rules, deliberately tuned somewhere between too many and too few "births" and similarly between too many and too few "deaths." It has an interesting near-equilibrium ecological statistics, with species like semaphores, blinkers, pinwheels, gliders, glider-guns, etc. This game developed a veritable world-wide cult, and so we are in possession of a large number of interesting constructions and observations. It is computation-universal, in the sense that one can assign to it an initial configuration that acts like a Universal Turing Machine (see mathworld.wolfram.com/GameofLife.html for details). What more shall we demand?

As long as we choose to use the game of Life as a Turing machine, see my response to the similar proposal for von Neumann's cellular automaton and that given in the Introduction. Universality notwithstanding, what a waste to use an entire two-dimensional parallel-processing universe for building a sparse and slow one-dimensional UTM where only one spot at a time is active (i.e., where the head is "now"). Building a UTM with Life was of course a clever tour-de-force as a way of proving the host environment's universality. But after that one should move on.

And here is where Fermi's question, "So where are they?" is appropriate. Tens of thousands of people must have striven to get some form of "life" emerge out of Life, and so far have failed. This is certainly circumstantial evidence that Conway's idea was stimulating—but the game of Life should not be regarded as a sacred relic to be held under a glass bell.

As for using the game of Life as an ecological universe for exploring the emergence of life, my impression is that it gets gummed up much too soon, as you will convince yourself if you look at a square of say, $10,000 \times 10,000$ cells started from a random initial condition and run it for 100,000,000 steps. Conway's Life is too jumpy on short-range interactions, and too refractory to long-distance interactions. One reason for that is that the non-invertibility of Life—the fact that it has orbits that *merge* (see [7] for a more thorough discussion), leads to a gradual "draining" of the effective space state, until the rule becomes effectively invertible and the orbits that are left are mostly short closed ones.

So we should not stop at that particular game of Life, but follow Conway's *spirit* and develop versions that have a well-argued promise for that "equilibrium near the edge of chaos" that life seems to thrive on. This rationale is explained in Sect. 4.

4 Specific Ergodicity

Possibly the parameter that most directly affects the capabilities of a distributed dynamical medium, like a cellular automaton, to support the emergence of complex structures, is the dependence of interaction strength on distance.

Complex, coordinated, knowledge-rich structures require memory—lots of it. And the characteristic property of memory is, of course, resistance to disturbances: I want to find in it what I put in it in the first place! Another important property of memory is, of course, ease of access, primarily reading access, but secondarily also writing access. In our DNA, which is vitally important, writing is done only at fabrication time; thereafter DNA is read-only. Moreover, DNA never leaves the tight shelves of the nucleus where it is stored. For work memory, the librarian will only allow you to come in person and copy a few pages at a time; you are free to use and reuse that, and throw it away when you're done. On the other hand, neural memory, which needs continual real-time upgrading, is stored at synapses, where it can be modified, though only gradually. Moreover, the same neuron firing is registered—even though with different interpretations—by thousand of synapses; so there is some redundancy there.

Other parts of an organism have to be constantly alert and active—a continual state of receptivity and change. But still you want to be selective about *what* is actually allowed to impinge on you, and your movements must be checked by feedback, lest you break everything you touch or you are broken by everything that touches you!

Organic life has managed to synthesize different materials wonderfully attuned for different functions—bone, skin, hair, mucus, hydrochloric acid, fat, crystalline lens, hemoglobin—and able to appropriately modulate and filter internal and external solicitations. On the other hand, a single general tuning knob ranging from REFRACTORY to HYPERACTIVE for the universe as a whole would be convenient for discovering the most effective MID-RANGE tuning. We have explored such a parameter, called *specific ergodicity*, in [6].

Basically, conventional ergodicity is a YES/NO parameter (a system is *ergodic* if all its states belong to a single orbit, or are in some probabilistic sense accessible from any other point). On the other hand, *specific ergodicity* is a parameter $0 \leq \eta \leq 1$, for distributed systems like cellular automata for which it makes sense to speak of an arbitrarily large wrapped-around patch governed by the same local rule, and to take the limit for the size of the patch going to infinity. In other words, for a situation where one can speak of a distributed system as a *material* rather than an *object*.

An invertible cellular automaton has $\eta = 0$ if every state has its own orbit of length 1, so that there is no interaction whatsoever between sites—every site is an isolated "monad"—and $\eta = 1$ if all states are on a single orbit (or a small number of orbits of almost maximal length), and thus in the long term the state of any site will be correlated with that of any other state—you can never be free of disturbances by neighbors no matter how distant.

A sample of well-known cellular automata gave values for η scattered over the whole range $[0, 1]$; thus, whatever this parameter means, the parameter appears to be *informative*—to be able to tell classes of systems apart. (If most of a library's books had a call number that began with P, the first letter of a book's call number would carry very little information—and so be of little use in classification.)

5 Emergence

Emergence is a term that scientists use in a specialized sense: very briefly, a pattern is *emergent* if it spontaneously arises (literally, "comes to the surface") from an underlying patternless substrate. Emergence is a class of statistical phenomena associated with a dynamical system.

A canonical example is the regular pattern of dunes that naturally develops out of a flat expanse of sand under the action of a steady breeze. Why do we get *ripples* if both sand and wind are *smooth* to begin with? And who insured the ripples' uniform spacing and specified its pitch? Based on animate beings' practices they were familiar with, the ancients fancied that the sun was pulled across the sky by a cosmic charioteer and lightning bolts were supplied to Zeus by a cosmic forger. Just as well, they might have fancied that dunes are raised by a cosmic potter tracing grooves with his fingers, or are turned up by the plowing of a cosmic farmer—techniques called respectively "fluting" and "furrowing."

But in fact, given the wind, the dunes arise by themselves with no external help, planning, or coordination. What happens if we come with a tractor and flatten the whole dune field? This has been done, and in a few days the dune field re-grew, with the same orientation and spacing—even though not necessarily with the same *phase*. (For example, the new ridges may have been arisen where in the previous pattern there occurred valleys instead.)

The most trivial form of emergence is the approach to thermal equilibrium in a homogeneous medium—the very phenomenon that led Fourier to invent his transform. When two blocks of copper at different temperatures are brought into intimate contact, the initially sharp step-function temperature profile along the resulting bar immediately starts smoothing into a sigmoid, which keeps stretching and flattening, until, in the limit of equilibrium, the temperature profile reduced to a constant function—a horizontal straight line. In Fourier's analysis, the short-wavelength components of the temperature plot decay fastest, until "every valley shall be filled, and every mountain … brought low," at all scales.

What's important is that this kind of macroscopic behavior happens all by itself—it need not be programmed in microscopic detail. This behavior is *robust*; even under perturbations, valleys will still be filled and mountains levelled. Molecules do not have to coordinate with one another. In fact, Fourier's analysis shows that they'd better not, because this particular emergent behavior is symptomatic of *linearity*, and thus of *noninteraction* between modes.

By its nature, the emergent phenomenon of thermal diffusion is *transient*—a one-act show. However, if we put a copper bar into contact at the two ends with two heat reservoirs at different temperatures, so that heat will be continually "pumped" into one end and "drained" from the other, then the temperature profile will converge to a *sloping* straight line. If we change the temperature difference, the slope will change. If we chill the block by splashing water on it, the temperature profile will be altered, but it will automatically revert to a sloping straight line soon after all the water has evaporated. Our temperature profile behaves like a stretched rubber band in a tub of molasses. You can pull it up and down and sideways, but when let go it will recover its equilibrium profile. In this case, the emergent state is *dynamically* maintained thanks to the flow of energy set up by the pump. Again, no micromanagement is needed for all of that, no exotic compounds or precisely machined part, no delicate assembly. It all just happens by itself, dependably resiliently, and predictably. If we may have a complaint, it is that what happens is only indirectly determined by us. If we get a sigmoid, we can't ask, "How about a sinusoid, instead? Emergence is more like a Chinese diner, where we can only order by number, than an *à la carte* Parisian restaurant. (Incidentally, Galileo was convinced that the shape of a rope hanging between two points was a *parabola*, which he was most familiar with; now we know it's a *catenoid*, very similar but not quite the same. It was not up to him to choose.)

Take now the flow of water through a pipe, or of a river between its banks. In either case the flow is maintained by a pressure difference, actively maintained by a pump in one case and the riverbed slope in the other.

As long as the water flows slowly, the velocity profile across the channel will be a smooth parabola, from a maximum at the center to zero at the edges; another nice mathematical shape spontaneously emerging from a compromise between friction between adjacent layers of water and friction between water and the edges of the channel—not a shape deliberately drawn by a cosmic draftsman.

Things get exciting when the pressure is increased and the flow runs faster. The layer of water running close to the bank develops lateral instabilities—it starts undulating. As the water speed increases, these undulations break up into vortices. Soon larger vortices develop smaller satellite vortices, and so forth, so that if you observe carefully you'll see a whole fractal hierarchy of them, spanning the range from macroscopic to microscopic. These vortices have an identity and a life of their own: they form and melt away, grow and shrink, they collide and annihilate one another, they "calve" daughter vortices, and so forth. They also may have "memory" and hysteresis: as you increase the speed, it may take a while for vortices to develop, but once they are there they will linger on even after you've reduced the speed to quite below what it was when they were born.

A similar phenomenon occurs with a dripping faucet, or when you overblow a note in a recorder. Also, the breakup of a thin stream of water into droplets may be encouraged by external vibrations, so that one can obtain in that way a very sensitive detector of ground temblors. What's more, as you adjust a faucet's flow, a train of regularly spaced droplets may break up into a train of paired droplet, from plic, plic, plic to plic plic, plic plic, plic plic; and this doubling may reduplicate, so that you may get a train like ((..)(..))((..)(..)). In some sense, this system has "learned" to count.

Here we've barely scratched the surface of emergence. It's an extremely rich and surprising world. On one hand, the ecologic "bestiary" supported by one substrate—say, the blinkers, gliders, and pinwheels of the game of Life—may be totally alien to those of another substrate—say, Norman Margolus's "critters" or Charles Bennett's "scarves" [4]. On the other hand, emergent phenomena are statistical phenomena, and naturally fall into a hierarchy that is basically one of combinatorial dynamical patterns. Diffusion, waves, cyclic rankings, frustration, annealing, predator-prey equation, hydrodynamics, single body "collisions" (like nuclear decay), two-body collisions, multi-body collisions, conservation on networks (like Kirchhoff's laws), relaxation oscillators, excitable systems, and so forth. Thus, even if the material substrate is different, system-level behaviors may obey similar laws and thus belong to the same equivalence class of phenomena.

There is an analogy here with the functions and the differential equations of physics—which, incidentally, are themselves mostly echos of emergent combinatorial phenomena. In principle, you could have all sorts of functions and differential equations, but in practice only a handful of them are truly common and ecumenical.

As we've seen, emergent behavior needs to be driven by a "pump," or what's called an "entropic cascade." (In the case of thermal equilibration between two bodies, the thermal difference only provides a single-use "battery" that ends up discharged when this difference reaches zero.) When an ordered arrangement of system A tumbles down the hill of increasing entropy turning into a more disordered arrangement A', some of A's "predictability" can be tapped and used to "pump up" another system from a more disordered state B' to a more ordered state B. An important case is when the peculiar order in A is of no special interest to us *per se*—all that we care is the *amount* of order that it contains. By coupling system A to system B we can extract some of the *generic* predictability of A and convert it into an equivalent amount of a *specific*—more desirable—form of predictability in B. (For example, if A is energetic but inedible wood that releases fire as it decays into A' (ashes and CO2), the fire can be used to bring a dish B' of raw meat or potatoes to a more edible state B.) As soon as the pump stops, the emergent structure "propped up" by it may of course collapse.

6 Life and Evolution

Life—and the opposite side of the coin, namely, evolution—is literally the *runaway daughter* of emergence.

Given a reliable entropic pump (this is really what is usually meant when one speaks of a "source of energy"), such as our Sun for the Earth, an emergent pattern based on it—artificial and propped up as it may be—constitutes an additional environmental niche available for habitation by materials, processes, and reactions, and may very well be one that provides tools and goods at a higher structural level than those provided by the base environment.

For example, even though the native language of a computer is its *machine* language, once an *assembler* for that machine language is available and can be run on that computer, one can program the same computer much more conveniently and productively at the assembler language level. In a similar way, C is a generic, platform agnostic, higher-level programming language. No computer runs C natively, but C *compilers* have been written to translate C to the machine language of different computers. And one can program more expressively, compactly, and productively in C.

Today. most of the Information Technology industry relies on different emergent levels of software environments, with enormous advantages in productivity, reliability, documentability, and often (though not always) in computer performance. Of course, all of today's civilization runs exclusively on the entropic pump provided by the sun either directly or via underground caches.

The runaway aspect of life mentioned above is that, starting about four billions years ago, life has managed to exploit natural emergent environments first, and then created new emergent environments within itself, originally at a sluggish rate, but then faster and faster (see Nick Lane outstanding books [1,2]).

In my opinion, the best way to study the emergence of life—beside the study of current and paleontological life on Earth—is to design environments representing versions of physics *stylized and domesticated* at different levels (we still don't know how much we can get away with) and optimized for the support of emergent environments. It is not terribly important at this point to create special-purpose hardware platforms for these environments; software platforms will do until we have a better idea of what we want.

An important issue in this context is to choose whether to stress (a) Non-invertible models of physics, such as von Neumann's, Bank's, Conway's, and Wolfram's cellular automata, whose merging of trajectories, as we've seen in Sect. 3.3 [Conway], provide a free, built-in entropy pump at least initially; (b) Models of physics which retain invertibility and provide an entropy pump through a distinguished away-from-equilibrium initial configuration (one containing the equivalent of a tame Big Bang, as it were); or (c) Models which replace an internal entropy pump by contact with an external thermal reservoir, as is routinely done to standardize thermodynamics arguments.

7 Conclusions

We all have learned a lot about modeling life within an automaton from the past two generations, but still can't get rid of Fermi's taunting question, "So, where are they?" I'm afraid that before trying to impose our wishes on an automaton we should more humbly inquire of it, "What is it that *you* would like to do?" and then try to build on that answer.

I'm reminded of the Flea Circuses that were still seen in country fairs as late as fifty years ago (have you ever seen Charles Chaplin's *Limelight*?), where an old man with a glass-topped tray-box hanging from his shoulders would show and illustrate to an audience of "children, soldiers, and servants" the acrobatics of fleas he kept in the box. The fleas clearly didn't "understand" his commands, but he somehow managed to anticipate the kind of things they'd more likely do. He knew them, he cared about them, he "understood" them. He would build his show on the flea-y things the fleas would naturally do. I'm sure he could have made a working computer out of jumping fleas, with the fleas still "thinking" that they were doing their natural flea-y things (and that's indeed the only things they could be doing) instead of being part of a computer. Just like the seals in Sea world or even the "computers" in Feynman's team. To use them well you have to know them well and treat them well.

References

1. Lane, N.: Life ascending: the ten great inventions of evolution, Norton (2010)
2. Lane, N.: The vital question: energy, evolution, and the origins of complex life, Norton (2015)
3. 't Hooft, G.: The cellular automaton interpretation of quantum mechanics, December 2015. arXiv:1405.1548v3
4. Toffoli, T., Margolus, N.: Cellular Automata Machines: A New Environment for Modeling. MIT Press (1987)
5. Toffoli, T., Margolus, N.: Invertible cellular automata: a review. Physica D **45**, 229–253 (1990)
6. Toffoli, T., Levitin, L.: Specific ergodicity: an informative indicator for invertible computational media. In: Computing Frontiers 2005. ACM (2005)
7. Toffoli, T., Capobianco, S., Mentrasti, P.: When–and how–can a cellular automaton be rewritten as a lattice gas? Theor. Comp. Sci. **403**, 71–88 (2008)

8. von Neumann, J.: The general and logical theory of automata. In: La, J. (ed.) Cerebral Mechanisms in Behavior–The Hixon Symposium, pp. 1–31. Wiley (1951)
9. von Neumann, J.: The Theory of Self-Reproducing Automata (Arthur Burks ed.). Illinois (1966)
10. http://en.wikipedia.org/wiki/Von_Neumann_universal_constructor
11. Watson, J., Crick, F.: A structure for deoxyribose nucleic acid. Nature **171**, 737–738 (1953)
12. Zamecnik, P., Keller, E.B.: J. Biol. Chem. **209**, 337–354 (1954)

http://www.springer.com/978-3-319-39299-8

Cellular Automata and Discrete Complex Systems
22nd IFIP WG 1.5 International Workshop, AUTOMATA 2016, Zurich, Switzerland, June 15-17, 2016, Proceedings
Cook, M.; Neary, T. (Eds.)
2016, XVI, 199 p. 77 illus., Softcover
ISBN: 978-3-319-39299-8