International Journal of Theoretical Physics, Vol. 42, No. 2, February 2003 (© 2003)

A Digital Perspective and the Quest for Substrate-Universal Behaviors

Tommaso Toffoli[1]

Received February 8, 2003

Should the Digital Perspective essays collected in these proceedings be viewed more seriously than attempts to play "fundamental theory"—which even in the hands of an Eddington was hard to tell from mythology and numerology?

We argue that a nonfrivolous aspect of this Digital Perspective is its heuristic capacity: to help us guess which aspects of our understanding of nature are more "universal," more robust, more likely to survive theoretical and experimental challenges. Behaviors that are *substrate-independent*—that can, for instance, thrive well on a digital support, even though they are traditionally imagined as taking place in a continuum—are especially promising candidates.

KEY WORDS: digital models of physics; inference; machinery of nature.

A common fallacy, exploited by pseudo-science fare such as *Chariots of the Gods*? (Von Daniken, 1987), goes something like this: "Five thousand years ago, without the help of iron tools or draught animals, the Moltechs raised on their sacred hill a 100-ton monolith. Nobody knows how they did it: *therefore* they must have been helped by God (or by ancient Egyptian science, aliens from Sirius, and so forth)."

Scientists constantly put themselves at risk of committing (or at least inviting) the converse fallacy, which runs: "My colleagues are intrigued by behavior *A*, observed in the real world. After much effort, I have come up with a mechanism that reproduces behavior *A*: *therefore* this must be the way God himself (Nature, Evolution, etc.) does it." You will find a whiff of that tendency in Wolfram's magnum opus (Wolfram, 2002) and in Fredkin's essay here (Fredkin, 2003) (see a recent review in *La Recherche* (Postel-Vinay, 2003), and, if you ignore the rather different role that I credit nature with, even in my own article in this issue (Toffoli, 2003). It's an open secret that, whatever its ostensible academic or utilitarian rationale, much scientific work is at bottom propelled by a personal curiosity to know *how the heck "God" manages to "do it*."

[1] Electrical and Computer Engineering, Boston University, Boston, Massachusetts.

0020-7748/03/0200-0147/0 © 2003 Plenum Publishing Corporation

The personal drive a Fredkin or a Wolfram may have for building their models is *their* business. But where does Digital Perspective leave the rest of us? Or, better, what does it *leave us with*?

It can leave us with the knowledge—this is what I argue here—that certain types of behavior are so *generic* (that is, relatively speaking, so easy to achieve) that *almost any substrate* will support them. Much as a machine is called computation-universal if it will support any kind of computation, thus behaviors that are supported by almost any kind of substrate may rightly be called *substrate-universal* behaviors.

Ernst Mach (1838–1916), Wilhelm Ostwald (1853–1932), Max Planck (1858–1947), and other "energeticists" were impressed by the elegance and the internal consistency of the thermodynamics edifice. One could debate whether mechanics was a surface aspect of electrodynamics (or vice versa), or whether Newton's law of gravitation was but the first term in the power expansion of some more subtle law. For all they knew and cared, that kind of physics was mere *phenomenology*, somewhat incidental and really not that interesting. "Planck's belief in the absolute validity of the second law made him not only reject Boltzmann's statistical version of thermodynamics but also doubt the atomic hypothesis on which it rested" (Kragh, 2000). To paraphrase Averroes' argument about religion vs. philosophy (Averroes [Ibn Rushd], 2001), atoms and other such animistic tokens may satisfy the simple ontological needs of the common people, but for the philosopher-physicist (read "energeticist") only more abstract principles, such as those expressed by equations like

$$\left(\frac{\partial E}{\partial N}\right)_{S,V} = -T\left(\frac{\partial S}{\partial N}\right)_{D,V}$$

can hold fundamental *physical* status.

Much as history eventually proved them wrong on many counts, with the above attitude Mach, Ostwald, and Planck were absolutely right—if we just drop the word "physical" from the last line. Those equations *are* fundamental precisely because they *do not* belong to physics per se. In fact, thermodynamics is a phenomenological, emergent aspect of statistical mechanics, and statistical mechanics, in turn, is mere combinatorial accounting—and thus mathematics and logic, not physics. This is well expressed by Jaynes' choice of words for the title of his book, *Probability Theory: The Logic of Science* (Jaynes, 2002).

To illustrate my assertion that certain physical behaviors are fundamental as behaviors precisely when they are *not* fundamentally physical I'll give two examples, one sublime and the other crude.

For the first example I'll stick with Planck's *idée fixe*. Much as the second principle of thermodynamics is associated with the idea of *physical irreversibility*, this principle turns out to be but a tautological rewording—in macroscopic terms—of the assumption that the underlying microscopic dynamics is *invertible*—and nothing else! This principle holds for *any* invertible dynamics; it is so general and fundamental precisely because it demands so little of physics. Once that little is granted, mathematics does the rest. In brief, mathematics says, “Show me an invertible system, and I’ll show you one where the second principle holds!” and physics can only obey.

For the cruder example, imagine that aliens are watching the Earth from Mars. Their telescopes can’t resolve better than one mile; thus, they can see the forest but not the individual trees. Now, suppose we parachute a million buffaloes in one spot in the middle of the Great Plains. The buffaloes will of course start roaming at random, indifferently in all directions; from a distance, the herd will diffuse isotropically and will closely approximate a two-dimensional gaussian splotch with a mean radius growing like $\sqrt{t}$. The “footprint” of the distribution (i.e., the area where there is a nonzero probability of having a buffalo) will be a circle of linearly increasing radius.

On another day we parachute a million buffaloes in the middle of Manhattan. This is the home of the well-known *Manhattan metric*: buffaloes can only proceed along streets and avenues, on a regular square grid, and thus will perform a *discrete random walk* in two-dimensions; the herd will spread in a two-dimensional multinomial distribution. Strictly speaking, the latter distribution bears witness to the anisotropy of the grid: for instance, its footprint will be a linearly growing *diamond* shape pointing NSEW rather than a circle. But, in the long run, significant differences from a circularly symmetric distribution will remain only where the probability of a buffalo is astronomically small to begin with, and thus, experimentally, very hard to tell from zero: virtually all of the herd will be *isotropically* distributed.[2]

Thus, although the microscopics of herd dynamics in Manhattan is very different from that in the Great Plains, the two distributions are virtually indistinguishable. Just by seeing a circular herd distribution a Martian cannot readily conclude that the buffaloes have been parachuted in the Great Plains, where they can indulge in continuous, isotropic roaming, since the discrete, anisotropic Manhattan dynamics happens to yield essentially identical results. By the same token, the Martians cannot conclude that the herd is in Manhattan—even though this hypothesis is simpler and yet sufficient to explain the phenomena.

What the Martians have lost is a mere geographical fact (Great Plains vs. Manhattan); what they’ve gained is a philosophically much more interesting insight: diffusive dynamics is a special, *highly-degenerate* dynamics, and for that reason can be relied on being *oblivious* to many details of the landscape.

[2] Observe that x appears with an exponent of *exactly* 2 in the expression for the gaussian distribution $g(x) = \exp(-x^2/2)$. Therefore, if we consider two independent gaussians, $g(x)$ and $g(y)$, and transform to polar coordinates, $r^2 = x^2 + y^2$, $\phi = \arctan(y/x)$, we see that the joint distribution $g(x) \cdot g(y) = \exp(-r^2/2)$ is independent of ϕ.

A similar argument could be made about *special relativity*. College physics professors may struggle to get into their students' heads that according to special relativity there are no absolute time and space, no preferred reference frame. What they should underline, instead, is that the *mathematical form* of special relativity is of such a generic (I am tempted to say "degenerate") nature that it could have emerged as easily from a fine-grained dynamics based on an underlying discrete grid (a cellular automaton or a lattice gas (Smith, 1994; Toffoli, 1989)) as from a continuous substrate. The specialness of relativity lies in its being such a *generic* category of behavior that it can easily be implemented *in a lot of different ways* (cf. (Toffoli, 2003)).

In conclusion, whenever Digital Perspective finds a way to reproduce a type of behavior it increases our confidence that that behavior is an *easy* one—that is, that it can be achieved in a lot of different ways ("If I can play this game, then almost anyone can!") Paradoxically, that makes us less, rather than more, certain about which is the way that nature actually uses.

We can snatch victory from the jaws of defeat by agreeing to redefine Nature (I'm dead serious here) as "The family of behaviors which are viable in an overwhelming majority of substrates." With that, the Digital Perspective becomes an excellent touchstone for telling not so much how Nature *does it*, but what Nature *is* to begin with.

ACKNOWLEDGMENTS

This research was funded in part by the Department of Energy under grant DOE-4097-5. I'm indebted to Lucas Ritz Toffoli for intelligent editing of the workshop's transcripts.

REFERENCES

Averroes [Ibn Rushd]. (2001). *An Authoritative Treatise and Exposition of the Convergence Which Exists Between the Religious Law and Philosophy* (ca. 1180). Translated by Ibrahim Najjar as *Faith and Reason in Islam: Averroes' Exposition of Religious Arguments*, Oneworld, Oxford.

Fredkin, E. (2003). "An introduction to Digital Philosophy." *International Journal of Theoretical Physics* **42**(2), 189–248.

Jaynes, E. (2003). *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge.

Kragh, H. (2000). Max Planck: The reluctant revolutionary. *Physics World*, December, p 31.

Postel-Vinay, O. (2003). Les nouveaux démiurges ("The new creators"). *La Recherche* **360**, 33–37.

Smith, M. (1994). *Cellular Automata Methods in Mathematical Physics*, PhD Thesis, *MIT/LCS/TR-615*, MIT.

Toffoli, T. (1989). Four topics in lattice gases: Ergodicity; Relativity; Information flow; and Rule compression for parallel lattice-gas machines, In *Discrete Kinetic Theory, Lattice Gas*

Dynamics and Foundations of Hydrodynamics, R. Monaco, ed., World Scientific, Singapore, pp. 343–354.
Toffoli, T. (2003). What is the Lagrangian counting?" *International Journal of Theoretical Physics* **42**(2), 363–381.
Von Daniken, E. (1987). *Chariots of the Gods*? Berkley Pub. Group.
Wolfram, S. (2002). *A New Kind of Science*, Wolfram Media, Champaign, Illinois.