

Physics and Computation

Computation and its Limits. By Paul Cockshott, Lewis M. Mackenzie, and Greg Michaelson, Oxford University Press, Oxford, UK, and New York, 2012, vi + 239 pages, \$63.00.

In the past few months, I have twice read newspaper reports of computer technology bumping up against fundamental physical constants. The first was an announcement that an international team, led by Michelle Simmons of the University of New South Wales, had constructed a transistor consisting of a single atom. The second was an offhand remark, in an article about the Knight Capital fiasco, that because of the speed of light, 1 foot per nanosecond, stock trading companies now need to locate their offices close to the trading floor. If your office is a mile from Wall Street, and you are running a 10-GHz machine, then

BOOK REVIEW

By Ernest Davis

by the time you receive the latest market news and respond, you have fallen 100,000 machine cycles behind your competition; in these days of high-speed trading, when gigabucks are made and lost in nanoseconds, you might as well send in your orders by foot messenger. One hundred seventy-five years ago, the electric telegraph made the world small; the electronic computer has made it large again. As it happens, the scales of the atom and elementary particles play large roles in the analyses in the book under review; the speed of light is a very minor player.

Computation and its Limits is partly a discussion of the limits that physics places on computation, partly a more general discussion of the relation between physics and computation. The most important and informative chapters of the book are the later ones. Chapter 5, a discussion of thermodynamics and computation, is centered around Rolf Landauer's (1961) argument about the entropy generated with each use of an AND or an OR gate in a computer. Chapter 6 is an extended discussion of the principles of quantum computing. Chapter 8 reviews and refutes a number of proposals that have been made for constructing computers that violate the Church–Turing thesis.

Chapters 5 and 6 contain extensive, self-contained introductions to thermodynamics and quantum theory, respectively. Unfortunately, like many self-contained introductions, these are largely unnecessary for readers who know the subject matter and largely over the heads of readers who do not. I found myself in the first category for the thermodynamics and in the second for the quantum mechanics, despite having taken a semester course in quantum mechanics a quarter century ago. Moreover, after the thermodynamic preliminaries of chapter 5, the actual presentation of Landauer's argument is unsatisfying. The argument, in brief, is that, because an AND gate takes two bits as input and produces one bit as output, one bit of information is lost; one bit of information corresponds to $k_B \cdot \log 2$ of thermodynamic entropy. Two obvious questions arise: First, how do you know that every physical implementation of a calculational bit has this thermodynamic equivalent? Second, since the inputs are still present at the input ports, in what sense are they lost? Neither question, particularly the second, is adequately discussed.

A much clearer presentation can be found in the paper "Physical Limits of Computing" by Michael Frank (*IEEE Computing in Science and Engineering*, May 2002), which is not cited here. That paper, incidentally, also discusses some physical limits of computation not included by Cockshott et al., such as the maximum density of information per unit volume. Proposals for getting around the Landauer limit by using reversible computations that do not lose information date back to work of C.H. Bennett in the early 1970s. Cockshott et al. discuss and reject one rather far-fetched design by Fredkin and Toffoli that achieved this aim using elastic collisions among billiard balls. They do not, however, discuss or even cite any of the more recent work in the area.

Despite the gaps in my understanding, I found the chapter on quantum computing the most interesting and informative in the book. In particular, the discussion of the various interpretations of quantum theory—the Copenhagen interpretation, the many-worlds interpretation, and so on—seemed to me remarkably well-written and fair-minded; the difficulties involved in each theory are clearly explained and acknowledged. The description of the general theory of quantum computing and of the methods that have been considered for its physical implementation is extensive and detailed. Curiously, though, there is little discussion of what has actually been accomplished. The implementations of Shor's algorithm, which to date have succeeded in factoring 15, are not mentioned. The D-Wave machine is mentioned in passing as "controversial," but few details are given about either the machine or the controversy.

Chapter 8 deals with a number of proposals for constructing machines that succeed in computing "non-computable" functions. Some of these were clearly not meant as practical suggestions. In the plan by Etesi and Némethi, for example, the solution to a halting problem would be learned by a computer an instant before it passed the Schwarzschild radius of a black hole. Beggs and Tucker propose a computer that uses infinitely long rigid rods. Others rely on mechanisms that carry out infinite-precision arithmetic on real-valued quantities. Cockshott et al. knock all of them down without breaking much of a sweat.

The chapter and the book end, however, in a serious misstatement. A machine that solved an uncomputable problem, the authors write, would be of no use because its answer could not be checked for correctness, and it would not be a general-purpose machine because it would not be able to solve its own halting problem; rather, it would be a special-purpose measuring device. None of this seems right to me.

Suppose that friendly space aliens arrive tomorrow in their UFO and make us a gift of a laptop that runs Java programs; the F11 key can solve the halting problem for any Java program that is inputless and that does not make a system call to the F11 key itself. This machine is perfectly consistent with Turing's proof; indeed, it is simply an oracle machine, much studied in computation theory. We are naturally skeptical, so we test it with a few hundred difficult programs for which we happen to know the answer—the search for a solution to Fermat's last theorem, say. (We might write a program that systematically searches through all tuples of integers $x, y, z, n > 2$, for a counterexample to Fermat's last theorem. Since Wiles's proof, we know that this program does not terminate; but that is not obvious just from an examination of the code.) The machine gives the right answer for all the test problems. Then it seems to me that (a) we are entitled to say that we "know" that this computer solves the halting problem, in the sense in which we "know" pretty much anything else; and (b) this is a general-purpose computer, in the same sense that an ordinary computer is a general-purpose computer. I don't expect that this will happen, but it is not logically impossible.

The remainder of the book seems to me less interesting, with substantial philosophizing of a kind not appealing to me. An early chapter begins with a discussion of Wigner's article "The Unreasonable Effectiveness of Mathematics in the Natural Sciences." On the second page of that discussion, "mathematics"

is reinterpreted as “computation,” a manoeuvre that in my opinion throws away Wigner’s point. To take Wigner’s first example, Newton’s law of gravity

$$m_i \cdot \frac{d^2 \bar{x}_i}{dt^2} = \sum_{j \neq i} \frac{Gm_i m_j}{|\bar{x}_j - \bar{x}_i|^2} \cdot \frac{\bar{x}_j - \bar{x}_i}{|\bar{x}_j - \bar{x}_i|}$$

is elegant mathematically but unpleasant computationally, particularly when $n > 2$. If the planets moved along paths composed of straight-line segments parallel to the coordinate axes with integer-valued vertices at constant reciprocal integer speeds, we would be entitled to admire how well adapted the laws of physics are for easy computation. My view is that Wigner’s article is best read as an expression of wonderment, rather than as a question with a useful or meaningful answer.

Finally, chapter 7 has to do with the real numbers. Cockshott et al. argue, obviously correctly, that infinite-precision real numbers are not physically meaningful. I do not know that anyone has ever claimed otherwise. They conclude that computer scientists should be skeptical about arguments that involve the existence of infinite-precision real numbers, such as Cantor’s proof that there are uncountably many real numbers.

This seems to me a bad direction to pursue. Mathematicians, computer scientists, and even physicists continue to use the real number system not because they think it is a true description of physical reality at the tiniest scale, or out of a foolish Platonism, or a sentimental attachment to the paradise that Cantor, Bolzano, Weierstrass, and others built for us; rather, once you learn how the system works, it is much easier to think about than finite-precision numbers. Anyone who doubts this is invited to compare the axioms of the real number system to the IEEE standard for floating-point arithmetic.

Readers with a serious interest in the relation between physics and computation will certainly find this book worth a look. The door is open, however, for someone to write a much better book on the subject.

Ernest Davis is a professor of computer science at the Courant Institute of Mathematical Sciences, NYU.