15 July 1996

PHYSICS LETTERS A

ELSEVIER

Physics Letters A 217 (1996) 188-193

The physical nature of information

Rolf Landauer [1]

IBM T.J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, USA

Received 9 May 1996
Communicated by V.M. Agranovich

Abstract

Information is inevitably tied to a physical representation and therefore to restrictions and possibilities related to the laws of physics and the parts available in the universe. Quantum mechanical superpositions of information bearing states can be used, and the real utility of that needs to be understood. Quantum parallelism in computation is one possibility and will be assessed pessimistically. The energy dissipation requirements of computation, of measurement and of the communications link are discussed. The insights gained from the analysis of computation has caused a reappraisal of the perceived wisdom in the other two fields. A concluding section speculates about the nature of the laws of physics, which are algorithms for the handling of information, and must be executable in our real physical universe.

1. Information is physical

Information is not a disembodied abstract entity; it is always tied to a physical representation. It is represented by engraving on a stone tablet, a spin, a charge, a hole in a punched card, a mark on paper, or some other equivalent. This ties the handling of information to all the possibilities and restrictions of our real physical word, its laws of physics and its storehouse of available parts.

This view was implicit in Szilard's discussion of Maxwell's demon [1]. Szilard's discussion, while a major milestone in the elucidation of the demon, was by no means an unambiguous resolution. The history of that can be found in Refs. [2,3]. The acceptance of the view, however, that information is a physical entity, has been slow. Penrose [4], for example, argues for the Platonic reality of mathematics, independent of any manipulation. He tells us "... devices can yield only approximations to a structure that has a deep and 'computer-independent' existence of its own." Indeed, our assertion that information is physical amounts to an assertion that mathematics and computer science are a part of physics. We cannot expect our colleagues in mathematics and in computer science to be cheerful about surrendering their independence. Mathematicians, in particular, have long assumed that mathematics was there first, and that physics needed that to describe the universe. We will, instead, ask for a more self-consistent framework in Sec. V.

P.W. Bridgman, recognized as Nobel laureate for his work in high pressure physics, published a remarkable paper [5] in 1934. That was his attempt to wrestle with the paradoxes of set theory. His solution: Mathematics must be confined to that which can be handled by a succession of unambiguous executable operations. Bridgman's paper is essentially a want ad for a Turing machine, which came a few years later. In a remarkable coincidence Bridgman even uses the word *program* for a succession of executable instructions. Bridgman

[1] E-Mail: landaue@watson.ibm.com.

PII S0375-9601(96)00453-7

did not go on to discuss the *physical* executability of the successive instructions, but that is the additional requirement we emphasize here [6,7].

2. Quantum information

When we first learned to count on our very classical and sticky little fingers we were misled into thinking about information as a classical entity. In the binary case that that requires either a *0* or *1* state. But nature allows a quantum mechanically coherent superposition of *0* and *1*. That is a degree of freedom appreciated only in recent years, and its impact still needs to be fully understood [8]. This new possibility has been investigated in three different scenarios so far; undoubtedly the list will grow.

One of these scenarios deals with *quantum teleportation* [9]. Here we use two previously prepared and correlated quantum objects; an EPR pair. One is shipped to the transmitting end, the other to the receiving end. An interaction between this prepared object at the transmitting end and the source object whose state is to be transmitted is used to generate a classical signal. At the receiving end this classical signal interacts with the prepared object located there to generate a copy of the teleported source state. The state of the source object is changed in the interaction that generates the classical signal. Quantum teleportation is a subject of serious conceptual interest, but it is not clear that it is a practical recipe for something that really needs to be done, and will not receive further discussion.

The most developed and obviously useful of the three scenarios is quantum cryptography [10], which has been demonstrated successfully in real systems, though its eventual realm of applicability is still uncertain. This application makes direct practical use of the uncertainty principle. A stream of quantum information cannot be examined by an eavesdropper without leaving a mark on the measured stream. A communication link, in contrast to computation, subjects each bit to limited handling, and this minimizes the need for each operation to do *exactly* what it is supposed to do. Furthermore, it is well known [11] that in a communications link (or memory) occasional rare errors are easily remedied by redundancy. Quantum cryptography, therefore, typifies a promising direction for the use of quantum information, in contrast to quantum parallelism, which will be discussed next.

Paul Benioff [12] first understood that a purely quantum mechanical time evolution can cause interacting bits (or spins) to change with time just as we would want them to do in a computer. David Deutsch later [13] realized that such a computer does not have to be confined to executing a single program, but can be following a quantum mechanically coherent superposition of different computational trajectories. At the end we can gain some kinds of information that depend on all of these parallel trajectories, much as the diffraction pattern in a two-slit experiment depends on both trajectories. Eventually Shor [14] showed that this form of parallelism provides tremendous gains for the factoring problem, finding the prime factors of a large number. That, in turn, is important for cryptography, and as a result quantum parallelism has gained widespread attention. We cite only some of the most elementary surveys [15].

3. Quantum parallelism: A return to analog computation

An analog computer can do much more per step than a digital computer. But an analog computer, in which a physical variable such as a voltage can take on any value within a permitted range, does not allow for easy error correction. Therefore, in the analog computer errors, due to unintentional imperfections in the machinery, build up quickly and the procedure can go through only a few successive steps before the errors accumulate prohibitively. A digital computer, by contrast, allows only a *0* or *1*. That permits us to restore signals toward their intended values, before they drift far away from that. In typical digital logic the signal is restored toward the power supply voltage or ground at every successive stage. This is what permits us to go through a tremendous number of successive digital steps, and this has given the digital computer its power. In quantum parallelism we do not just use *0* and *1*, but all their possible coherent superpositions. This continuum range, which gives quantum parallelism its power, also gives it the problems of analog computation, a point first explicitly stated by Peres [16]. If we have a state which is mostly a *1*, with a small admixture of *0*, we cannot simply eliminate that admixture;

the superposition may be the intended signal.

Imperfections in the quantum mechanically coherent computer generate several separate problems. First of all, interaction of the intentional information bearing degrees of freedom with the environment causes decoherence and spoils the quantum mechanical interference between alternative trajectories which constitute the basis of quantum parallelism. In the literature generated by the advocates this is the most widely recognized of the several problems faced by quantum parallelism. We can give only a small sampling of that recognition [17]. Note that in order to carry out logic the information bearing degrees of freedom must interact strongly with each other. At the same time, to preserve coherence, they cannot interact with anything else, including the physical framework that holds the bits in their positions. That is a tall order! A second problem arises from the fact that there are manufacturing flaws; the machinery will not do exactly what is intended. For example, interacting bits may not be spaced exactly as needed. Or the external radiative pulses, invoked in a number of the more detailed embodiments, may not meet their exact specification. A $\pi/2$ pulse may be a little too long, too short, or come too early. Flaws in the machinery can cause two separate problems. First of all they can, and generally will, introduce erroneous components into the states reached by the computation. Additionally, however, they can cause unintended reversals of the computational process; they can reflect the motion of the computer. This problem exists most clearly in the case of the Feynman computer [18]. In this computer the computation is launched as a wave-packet moving down its computational track, much as an electron can be sent down a one-dimensional periodic chain of atoms. The one-dimensional electron case is beset by *localization*: transmission diminishes exponentially with length due to irregularities in the supposedly periodic potential. We can expect the same for the computational trajectory. Most of the recent quantum-mechanical computer proposals are not Feynman computers; they do not propagate with their initial kinetic energy. Rather, they are clocked externally. One can then *hope* that if we push the computation forward in a sufficiently determined manner (strong arm and stiff crank?) reflections can be avoided. That is a hope, not yet backed by analysis. One clocked scheme using time-dependent Hamiltonians has been analyzed [19], and does exhibit a reversal problem. *Localization* is a condensed matter theory concept, whereas quantum computation has been studied by computer scientists and by physicists who (largely) do not have a condensed matter background. That is why unintended reversals are not even discussed by the advocates. For exceptions see Ref. [20].

Error recognition and error correction in quantum computation cannot follow the recipes we learned for classical digital computers. Error recognition requires the ability to distinguish a signal from its ideal value. But we cannot, in general, tell whether two arbitrary quantum states differ, or not. Even if we were able to recognize errors, we cannot throw away the description of the error. Discarding information is a dissipative event and will spoil the coherence needed for quantum parallelism. If we do keep a record of the error it must be led aside, so as not appear in the subsequent interference. Despite these difficulties progress has been made toward error reduction, and we can cite only a sample of the material on its way [21]. This is far more progress in fact than this author thought possible, but not enough to permit computation. An effective error correction approach must work for *all* logic steps, and must not rely on perfect supplementary apparatus nor on additional signals which have to be presumed to be perfect. Undoubtedly, further progress will be made, but victory is not yet in sight.

A particularly serious difficulty is caused by components which, even if they meet the ideal specification, still do not do exactly what they are supposed to do. Ref. [19] listed the difficulties faced by Lloyd's proposal [22]. Ref. [23] listed the subtle deviations from those that are actually required in the steps needed for a quantum communications scheme. We suspect that most proposals, if analyzed equally carefully, will show similar flaws. If the devices, even in their ideal state, do not do *exactly* what is needed, that becomes a particularly stubborn problem. At best, only interaction of the results with those obtained in a totally different way can give any hope of error reduction, and it is not clear how this can be achieved and whether it can provide adequate correction.

Finally there is a computer science problem, unrelated to the physics and technology we have stressed. Even if quantum parallelism can be made to work, what is its range of application? The world is unlikely to want to pay for the development of a difficult technology without a broad payoff. The suggestion has been made [24] that quantum computers are particularly good at simulating quantum systems. But for that to be true the system, which needs to be analyzed, must have a structure which maps easily and cleanly onto that of the computer. Would a one-dimensional chain of interacting two level systems [22] make it easy to follow electronic motion in a heavy atom?

Despite the pessimism I have expressed, the mere possibility of quantum parallelism has changed theoretical computer science permanently. Those concerned with theorems about the minimal number of steps required for the execution of an algorithm must allow for quantum parallelism.

4. Energy dissipation requirements

The history that led to our understanding of the energy requirements of the computer has been summarized elsewhere [2]. Reversible classical computation, in the presence of viscous friction proportional to velocity, can be accomplished with as little energy dissipation, per step, as desired. Occasional objections still appear to reversible computation and/or the need to dissipate energy when discarding information. Some of the objections are discussed and put aside by Shizume [26]. But most of the interested community long ago accepted these notions. It is, however, not widely recognized that the same body of work that led to reversible computing also demonstrated that any desired immunity to noise can be obtained, and that this is not limited to thermal noise. Similarly, in the measurement process, the reversible operations can be carried out with arbitrarily little dissipation. Resetting the meter, after it has become separated from the system to which it was coupled for measurement, is irreversible, and is the essential lossy step [3]. It is a sociological puzzle why this insight, achieved so long after Maxwell first posed the demon question, has not been celebrated more widely. As in the case of most scientific insights, a number of investigators contributed along the way as discussed in Ref. [27]; nevertheless, the clear, confident and complete resolution stems from the 1980's.

A comparable understanding of the energy needs of the communications channel came later. For many decades the general perception has been that it takes $kT \ln 2$ to send a bit in a classical channel, and more if $h\nu > kT$, where ν is a typical signalling frequency. This mode of thought has been described in more detail elsewhere [23,28]. It is based on the assumption that a linear boson channel is used and that the energy used in the transmission has to be dissipated. As stressed in Ref. [25], this whole field of concern with the physical fundamentals of information handling is characterized by first answers which have been wrong. Alternative communication methods have been proposed which, in the classical case [28,29] take arbitrarily little energy dissipation, if the bits are moved slowly. Slow motion does not imply a low bit rate; the bits can follow each other as densely as desired. Furthermore, as in the case of reversible computation, the classical communication schemes allow for any desired immunity to noise by suitable choice of the device parameters. The methods do not require perfect components. The classical conclusion should have been obvious once reversible computation is accepted. After all bits are moved around in a computer and therefore there can be no minimal unavoidable energy penalty for bit motion, if there is none for the overall computation. The quantum communication link does not allow an equally easy chain of reasoning. The existing theoretical treatments of the quantum computer allow only for the interacting information bearing degrees of freedom, and ignore all others degrees of freedom including those in internal communication links. Therefore the fact that the literature describes Hamiltonian quantum computers provides no insight into their internal communication links.

It has been shown, however, that if we optimistically assume that frictionless and quantum mechanically coherent machinery is available, then a nondissipative quantum communications link can be constructed [23]. It does depend, not surprisingly, on a sequence of operations each of which is reversible. The classical and quantum communication links proposed by this author are not practical recipes, but are only existence theorems. In particular they invoke active machinery all along the link, in contrast to the passive nature of an electromagnetic transmission line or an optical fiber. It is not clear, however, that this is an essential restriction [23]; there may be room for invention. Again there is a sociological mystery. When I first proposed low energy communication links, in the classical case, I did not expect ready acceptance for my refutation of the conventional wisdom. I expected angry rebuttal. In actual fact, with the exception of one dissenting publication [30], concerned really with reversible computation rather than the communications link, there was remarkable silence.

5. Impact on the laws of physics

This is by far the most speculative part of this discussion, but perhaps also its most significant aspect. I return here to notions first presented almost thirty years ago [6] and elaborated on a number of occasions after that, e.g. in Ref. [31]. The laws of physics are essentially algorithms for calculation. These algorithms are significant only to the extent that they are executable in our real physical world. Our usual laws of physics depend on the mathematician's real number system. With that comes the presumption that given any accuracy requirement, there exist a number of calculational steps, which if executed, will satisfy that accuracy requirement. But the real world is unlikely to supply us with unlimited memory or unlimited Turing machine tapes. Therefore, continuum mathematics is not executable, and physical laws which invoke that can not really be satisfactory. They are references to illusionary procedures. Can we not prove that $\cos^2\theta + \sin^2\theta = 1$ exactly, and not just to some very large number of bits? Yes, within a closed postulate system that can be demonstrated with a limited number of steps. But the laws of physics need to go beyond that and require actual number crunching. In a world with limited memory we simply cannot distinguish between π and a terribly close neighbor.

In fact, the limit on memory size is not the only limitation likely to be imposed by the real world. Computer elements are continually beset by an endless variety of adverse influences. Cosmic rays, alpha particles, electromigration, corrosion, spilled cups of coffee and earthquakes are a partial list. Can we keep going without limit in the presence of this exposure? Can we invent hardware which can be made arbitrarily immune to degradation? If we simply make the equipment more massive, then we will exhaust the universe's supply of available parts more quickly. At this point the skeptic will respond: Our inability to calculate the exact evolution under a law of physics does not prevent the universe from following that law. The evolution is precisely determined; it is only our ability to simulate that on a computer which is limited. The answer to that: It is not an incorrect response; it is an unverifiable and therefore meaningless response. The skeptic of my proposed statement will suggest that we can divide the universe into two halves, set them into identical initial states lending to the same subsequent history. Quantum mechanics does not, of course, allow us to be sure we have set two systems into the same state [32]. But even if we ignore that, how can we possibly be sure that two such large entities, requiring remarkably complex description, are alike? There cannot be a meaningful wave function for the universe, and very likely not even for half of it.

The limited precision available for the execution of algorithms implies that there is some sort of uncertainty in the laws of physics. This is not necessarily exhibited in the form of a limit to a certain number of bits, it may be more stochastic in nature. This may in turn be the ultimate source of noise and of irreversibility in the universe [28,33].

Even if the speculation just presented is not acceptable to the reader, our argument that mathematics, or any information handling process, cannot reasonably invoke an unlimited number of successive steps, is broader and more compelling. We cannot, of course, expect mathematicians to give up their beautiful array of concepts without hesitation and struggle.

A reluctance to accept the continuum can be found in many places. See for example John Wheelers's frequent slogan "no continuum" [34], and his citation of others with that view [35]. John Wheeler also has a second and more significant relation to this discussion. Our scientific culture normally views the laws of physics as predating the actual physical universe. The laws are considered to be like a control program in a modern chemical plant; the plant is turned on after the program is installed. *In the beginning was the Word and the Word was with God, and the Word was God* (John I,1), attests to this belief. *Word* is a translation from the Greek *Logos* "thought of as constituting the controlling principle of the universe" [36]. Wheeler, in a number of discussion [37] has an adventurous view in which the laws of physics result from our observation of the universe. Wheeler's details are not my details, but we both depart from the notion that the laws were there at the beginning. The view I have expounded here makes the laws of physics dependent upon the apparatus and kinetics available in our universe, and that kinetics in turn depends on the laws of physics. Thus, this is a want ad for a self-consistent theory.

References

[1] L. Szilard, Z. Phys. 53 (1929) 840.
[2] C.H. Bennett, IBM J. Res. Dev. 32 (1988) 16.
[3] H.S. Leff and A.F. Rex, Maxwell's demon: entropy, information, computing (Princeton Univ. Press, Princeton, 1990).
[4] R. Penrose, The emperor's new mind (Oxford University Press, Oxford, 1989).
[5] P.W. Bridgman, Scr. Math. 2 (1934) 3.
[6] R. Landauer, IEEE Spectrum 4(9) (1967) 105; Reprinted with added commentary in Speculations in Science and Technology 10 (1987) 292.
[7] A. Ekert, in: Proceedings of the 14th ICAP, eds. D. Wineland et al. (AIP, New York, 1995) p. 450.
[8] C.H. Bennett, Phys. Today 48 (October 1995) 24.
[9] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W.K. Wootters, Phys. Rev. Lett. 70 (1993) 1895.
[10] I. Peterson, Science News 149 (10 February 1996) 90; A. Muller, H. Zbinden and N. Gisin, Europhys. Lett. 33 (1996) 335.
[11] C.E. Shannon, Bell Syst. Tech. J. 27 (1948) 379, 623.
[12] P. Benioff, J. Stat. Phys. 22 (1980) 563; J. Stat. Phys. 29 (1982) 515; Phys. Rev. Lett. 48 (1982) 1581.
[13] D. Deutsch, Proc. R. Soc. Lond. A 400 (1985) 97.
[14] P.W. Shor, in: Proc. 35th Annual Symposium on the Foundation of computer science (IEEE Computer Society, Los Alamitos, CA, 1994) p. 124.
[15] S. Lloyd, Sci. Am. 273 (1995) 140; D. DiVincenzo, Science 270 (1995) 255; J. Glanz, Science 269 (1995) 28; B. Schwarzschild, Phys. Today 49 (1996) 21.
[16] A. Peres, Phys. Rev. A 32 (1985) 3266.
[17] I.L. Chuang, R. Laflamme, P.W. Shor and W.H. Zurek, Science 270 (1995) 1633; M.B. Plenio and P.L. Knight, Phys. Rev. A 53 (1996) 2986; W.G. Unruh, Phys. Rev. A 51 (1995) 992.
[18] R. Feynman, Opt. News 11 (2) (1985) 11; reprinted in: Found. Phys. 16 (1986) 507.
[19] R. Landauer, in: Proc. Drexel-4 Symposium on Quantum nonintegrability – quantum-classical correspondence, eds. D.H. Feng and B-L. Hu (International Press, Boston), in press.
[20] S.C. Kak, Found. Phys. 26 (1996) 127; P.B. Benioff, Hamiltonian models of quantum computers which evolve quantum ballistically, preprint.
[21] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin and W.K. Wootters, Mixed state entanglement and quantum error-correction, Phys. Rev. A., to be published; R. Laflamme, C. Miquel, J.-P. Paz and W.H. Zurek, Perfect quantum error correction code, to be published; P. Shor, Phys. Rev. A. 52 (1995) 2493.
[22] S. Lloyd, Science 261 (1993) 1569.
[23] R. Landauer, Minimal energy requirements in communication, Science, to be published.
[24] R.P. Feynman, Int. J. Theor. Phys. 21 (1982) 467; S. Lloyd, Universal quantum simulators, preprint.
[25] R. Landauer, in: Proc. Workshop on Physics and computation PhysComp '94, (IEEE Computer Society Press, Los Alamitos, 1994) p. 54.
[26] K. Shizume, Phys. Rev. E 52 (1995) 3495.
[27] R. Landauer, Physica A 194 (1993) 551.
[28] R. Landauer, in: Selected topics in signal processing, ed. S. Haykin (Prentice Hall, Englewood Cliffs, 1989) p. 18 (printed version has some figures oriented incorrectly).
[29] R. Landauer, Appl. Phys. Lett. 51 (1987) 2056.
[30] W. Porod, Appl. Phys. Lett. 52 (1988) 2191.
[31] R. Landauer, in: Proc. 3rd Int. Symp. on Foundations of quantum mechanics – in the light of new technology, eds. S. Kobayashi, H. Ezawa, Y. Murayama and S. Nomura (The Physical Society of Japan, Tokyo, Japan, 1990) p. 407.
[32] W.K. Wootters and W.H. Zurek, Nature 299 (1982) 802.
[33] R. Landauer, Z. Phys. B Condensed Matter 68 (1987) 217.
[34] J.A. Wheeler, in: Physical origins of time asymmetry, eds. J.J. Halliwell, J. Perez-Mercader and W.H. Zurek (Cambridge Univ. Press, Cambridge, UK, 1994).
[35] J.A. Wheeler, IBM J. Res. Dev. 32 (1988).
[36] Webster's New World Dictionary of the American Language, ed. D.B. Guralnik (William Collins + World Publishing, 1978).
[37] J.A. Wheeler, in: Problems in theoretical physics, A. Giovanni, F. Mancini and M. Marinaro, Eds. (University of Salerno Press, Salerno, 1984) p. 121; in: Frontiers of nonequilibrium statistical physics, eds. G.T. Moore and M.O. Scully (Plenum, New York, 1986).