ZIG-ZAG PATH TO UNDERSTANDING

ROLF LANDAUER
IBM Thomas J. Watson Research Center
Yorktown Heights, NY 10598

ABSTRACT

Our understanding of the fundamental physical limits of information handling has developed along a very convoluted path. Most of the initially plausible physical conjectures have turned out to be wrong. A participant's personal view of these events is not a disciplined contribution to the history of science. I do, however, list my own mistakes along with those of others.

1: Computation

The attempt to understand the ultimate physical limits of the computational process is almost as old as the modern electronic digital computer. This search was stimulated by the earlier examples of thermodynamics which arose from the attempt to understand the limits of steam engines, and by Shannon's channel capacity theory. These examples suggested that the computer could be characterized in a similar way. The earlier prototypes had derived optimum performance limits in a way which transcended particular technologies and design choices. Scientists and engineers take pride in the ability to do back–of–the–envelope calculations, to quickly reach to the critical aspects without encumbering details. Yet, in this field of ultimate physical limits of information handling, the quick and dirty approaches have turned out to be wrong in a remarkably consistent way. All of the first answers have been misleading.

In the 1950's it was natural to associate a binary degree of freedom with kT and to assume that a minimal energy of that order had to be associated with an elementary logic step, to provide noise immunity. The identification of "associate with" with required energy dissipation tended to be made in casual blackboard discussions. It was not until 1961 [1] that it became clear that the process which really required a minimal and unavoidable energy penalty was the discarding of information.

The assumption that an energy dissipation of the order of kT was required by every logic step seemed, in the 1950's, to be a natural consequence of "known" results. It was "known," after all, that it took $kT\ln2$ to send a bit along a communications line, and computation required the frequent transmission of signals. It was also "known" that measurement required energy dissipation. Unfortunately, as we now know, the communication results were really more limited than generally presumed. Furthermore the sophisticated literature on the measurement process often left it unclear what a measurement really is and how to tell a measurement from an elephant's trunk. Additionally, as is now known, classical measurement theory as then accepted suffered from its own blemishes which were not widely understood until the 1980's.

This subject, during the 1950's was mostly one for casual unpublished speculations. Ref. [2] tells us, for example, that " ...von Neumann [3] and Brillouin [4] conjectured that $kT\ln2$ minimum energy must be spent for each step of information processing". Many other authors have related statements, e.g. Igeta's reference to Brillouin [5]. Brillouin's famous book [4], despite a chapter *The Problem of Computing,* does not allude to the actual logic processes in a computer, e.g. to a logical *and* or a logical *or*, and contains no references to a total working computer, such as a Turing machine or a cellular automaton. Arthur Burks [3], nine years after von Neumann's death, and five years after Ref. [1], credits von Neumann with the notion that a computer must dissipate at least $kT\ln2$ at room temperature "per elementary act of information, that is per elementary decision of a two–way alternative and per elementary transmittal of one unit of information." This comes closer to a sensible discussion than can be found in Ref. [4], but still involves the ambiguous *per elementary decision of a two-way alternative.* Indeed, we now know [6], the mere transmission of information does not require any minimal energy dissipation [7].

von Neumann's insight into the computational process and associated physics was impressive. One of his patents [8], assigned to my company, has exerted a tremendous influence on my own work. I do not

0-8186-6715-X/94 $4.00 © 1994 IEEE

question that von Neumann *could* have answered the energy dissipation question properly, and perhaps did. But the written record demonstrating that has not been published.

In 1961, when Ref. [1] appeared, it was natural to assume that discarding information was an inevitable part of computation. After all, computers did lots of that. Ref. [1] already appreciated that information discarding operations could be imbedded in larger operations which were 1:1. But that did not constitute an understanding of reversible computation; that had to await Charles Bennett's paper [9] which Wheeler and Zurek have labelled *epoch–making* [10]. In particular, I had assumed (quite incorrectly), that a computation which runs along a chain of 1:1 transformations is a table look–up device, where the designer has to anticipate every possible computational trajectory. Bennett's insight that computation can utilize a series of steps, each logically reversible, and that this in turn allows physical reversibility, was counter–intuitive when it first appeared. (Reversible computation, without understanding of its physical significance, was first described by Lecerf in 1963 [11].) That computation could be done with arbitrarily little dissipation, per step, was not actually in contradiction to the principal thesis of Ref. [1], but seemed to go remarkably beyond that. It also distinguished the computational process in a surprising way from our perceptions, at that time, about communications and measurement.

The notions expressed by Charles Bennett and myself have been elaborated in many ways by others. Despite that, some grumbling persists as evidenced by two papers in the 1992 version of this workshop [5, 12]. This is best characterized by the excerpts from Ref. [13], based in turn on my own unpublished comments in the electronic forum resulting from the 1992 version of this workshop.

Nothing in science is ever settled totally, and beyond all question. In 1994 we will again see some challenges to the second law, and some proposals for superluminal signal propagation. A few of these may be genuine open–minded and scholarly attempts to probe the limits of our certainty. A larger number will reflect honest attempts to respond to a poor exposition with the challenger unaware of the diversity of paths that has led to the established conclusions. But the large majority...

... I did not believe or understand Charles [Bennett] *when he first explained his emerging notions to me, in 1971. It took me some months to come to understanding and agreement.*

It is no longer 1961 or 1973. Our concepts have been explored, expanded, and reformulated by many colleagues with different backgrounds and differing motivations. I, therefore, believe that our results are established as well as most scientific results can be, even though there will continue to be a modest flow of objections....

For a number of years all reversible computer embodiments were either impractical mechanical machinery, or else were based on nucleic–acid inspired chemical reaction sequences. I found myself searching for deep reasons for that; why were there no electrical versions? That was solved eventually, when Likharev [14] came out with a Josephson junction version. There was no deep mystery! On the other hand, Likharev's proposal was very demanding on the allowed variation of components. Furthermore it was tied to a technology which was about to disappear from the list of serious candidates.

The 1992 version of this meeting showed through realistic CMOS circuit proposals, that reversible computation was not just a tool for answering conceptual questions about limits, but a tool for reducing power consumption in reality [15]. Once again an unexpected turn, even though partially anticipated by the earlier work of Seitz [16]. Normally, in CMOS circuits, the energy stored in capacitances is discharged and dissipated in each switching event. In the proposals of Ref. [15] the charging process is reversed, and the energy taken back into (the suitably designed) power supply. The forward propagation of the signal, between latches, is simply reversed. Whether the added delays and complexity that come with these proposals leave them advantageous is a question to be settled by further exploration. These proposals, in turn, have inspired a version [17] which avoids the need for detailed reversal, but still saves most of the power.

Why was the possibility of reversible CMOS logic appreciated at such a late stage? In part perhaps, because the people interested in reversible computation were largely computer scientists or basically oriented physicists, rather than circuit or device technologists. But there were additional errors in perception, at least in my own case, and I suspect that others shared my mistake, which will be explained. Reversible computation as invented and described by Bennett [9], requires every single step to be reversible. It is not enough to save a copy of the initial state of the computer. Now an *and* or *or*, for example, are not 1:1 operations; the input cannot be deduced from the output, and therefore one might (and I did) assume that such operations cannot be utilized in reversible computing. Reversible computation is as shown in Fig. 1a, in contrast to

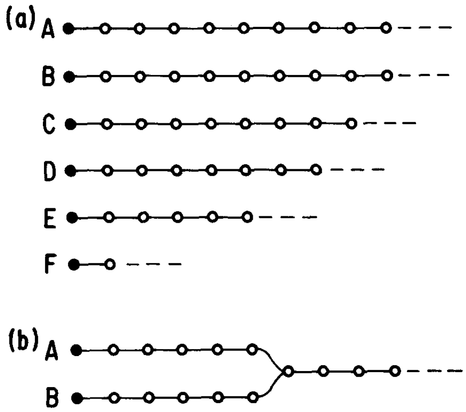

Figure 1: *a.* One–to–one computation. The left–hand end of a horizontal chain represents the initial state, and forward computation represents motion to the right, through a sequence of states represented by successive open circles. Different capital letters correspond to different initial states, that is, different programs. *b.* Information–discarding junction. Two computational paths, moving to the right, merge into one.

Fig. 1b where trajectories merge.

Now, however if, as in *Hansel and Gretel* you leave markers behind to identify your trail, Fig. 1b *is* reversible. You only need to spot and pick up the markers on the return trip. In the recent reversible CMOS proposals the markers are replaced by the input signal, which stays there. It remains there until the output reversal is completed. Thus, the usual simple logic circuitry with *and* and *or* gates can be used. Of course, if we want to save the capacitive energy in the latches it gets more complex; that will not be discussed here.

2: Measurement and communication

The historical zig–zag pattern also applies to two auxiliary fields: the energy required for measurement and the energy required for the communications channel. The history of Maxwell's demon which caused us to focus on the energy needed in the classical measurement, has been documented in detail in Refs. [13, 18]. Maxwell's demon utilizes the fact that if we have information about the motion of individual molecules, we can extract their kinetic energy to do useful work. Thus, we must make measurements on individual molecules. For a good many years it was assumed that the transfer of information from the object to be measured, to the meter, required a minimal and unavoidable energy dissipation, and that this dissipation saved the second law. Instead it is the resetting of the meter, to a standardized state, after each use (and before the first use), that requires the energy dissipation needed to save the second law. An adapted quotation from Ref. [19] describes this setting: *Well known discussions, by Brillouin [4], followed by a refined version due to Gabor [20], invoke the fact that to "see" a molecule a photon must be used. In order to distinguish it from the surrounding black body radiation, the photon must have an energy $h\nu > kT$. This energy is assumed to be lost in the process. A great deal of later literature of which we cite only a few examples, echoes these notions [21]. In retrospect, this acceptance of the Brillouin–Gabor view appears as one of the great puzzles in the sociology of science. If someone proposes a method for executing the "measurement,"... which consumes a certain amount of energy, why should we believe that the suggested method represents an optimum?*

The communications channel represents another similar episode. There is a widespread presumption that it takes $kTln2$ to send a bit from one place to another. It is implied that this follows directly from Shannon's results for the linear transmission line

$$C = W log_2 \frac{P+N}{N} \tag{1}$$

where W is the line's bandwidth, P the signal power, and $N = kTW$ the thermal equilibrium noise power. Shannon never claimed such a universal applicability for his result. I have supplied a number of counter–examples to show that a bit can be transmitted with arbitrarily little dissipation, if we are willing to do it slowly [19, 22]. Despite the fact that my analysis ran contrary to the prevailing wisdom, it has been ignored. Only Porod [23] has paid me the compliment of a debate, and Ref. [24] correctly characterizes the limits of the widely accepted viewpoint.

3: Quantum computers

Finally let me allude to the totally quantum mechanical computational process. I had attempted, in unpublished work, to describe such a process, e.g. quantum versions of the Bennett–Fredkin–Turing machine [25], but got hopelessly bogged down in the complexity of that. Eventually Benioff saw the way to do that [26]. You invoke a Hamiltonian (or a unitary time evolution) which causes the information bearing degrees of freedom to interact, and to evolve with time, as they do in a computer. You introduce no other parts or degrees of freedom. I was too engineering oriented to see that possibility; I assumed that you had to describe the apparatus and not just the Hamiltonian. Benioff's idealization was of course just that. A penalty paid to permit a theory. Benioff's work led to a wide literature which I have assessed elsewhere [6, 7, 27]. Feynman was a particularly significant contributor to the stream [28] that followed Benioff. Feynman was present at the 1981 workshop at M.I.T. [29] where many of us discussed Benioff's notions and the paper Benioff presented there [30]. Did we understand and believe Benioff? Feynman did not need much of a clue, and as a result generated his own very appealing and effective view [28] of quantum mechanical computation. Unfortunately, Feynman failed to cite Benioff. This meeting also provided Feynman's first exposure to Bennett's notion of reversible computation. Feynman understood immediately, an impressive feat in 1981. For an alternative and complementary view of Feynman's role in this field see Ref. [31].

Recently there has been a good deal of excitement about quantum parallelism, undoubtedly reflected in some companion papers in this volume, as well as in the earlier 1992 proceedings [32]. I cannot but help contrast that field to the discussions of the energy requirements of the communications channel. The proponents of quantum parallelism have provided us with a description of the computational Hamiltonian, not of apparatus. They have not, up to now, inquired about the consequences of Hamiltonians which deviate somewhat from the exact desired value. This generosity of spirit contrasts strikingly with that found in the communications channel literature. The latter concentrates on the linear boson channel, despite the fact that written communication, as well as the shipment of floppy disks, are well established ways of sending messages. Levitin's contribution [33] to the earlier 1992 workshop is typical of this literature.

4: Overview

The path to understanding in science is often difficult. If it were otherwise, we would not be needed. This field, however, seems to have suffered from an unusually convoluted path.

Conference records [29, 34] demonstrate that a good many perceptive investigators enter a field without any attempt to read the existing literature. That is a contributing factor to the difficulty we have noted. The lack of experimental data also permits us to stick to erroneous concepts. But these are auxiliary sources; the zig-zag path, ultimately, arises from the difficulty of the problem when compared to our ability.

References

[1] R. Landauer, "Irreversibility and Heat Generation in the Computing Process," *IBM J. Res. Dev.* vol. 5, 1961, pp. 183-191. [Reprinted in: *Maxwell's Demon.* H.S. Leff and A.F. Rex, eds. Princeton: Princeton Univ. Press, 1990, pp. 188-196].

[2] E. Goto, W. Hioe and M. Hosoya, "Physical Limits to Quantum Flux Parametron Operation," *Physica C* vol. 185–189, 1991, pp. 385–390.

[3] J. von Neumann, *Theory of Self-Reproducing Automata*, A.W. Burks, ed. Urbana: Univ. Illinois Press, 1966.

[4] L. Brillouin, *Science and Information Theory.* New York: Academic Press, 1956.

[5] K. Igeta, "Physical Meaning of Computation," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 184–191.

[6] R. Landauer, "Information is Physical," *Phys. Today* vol. 44, May 1991, pp. 23–29.

[7] R. Landauer, "Information is Physical," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 1–4.

[8] J. von Neumann, "Non-linear Capacitance or Inductance Switching, Amplifying and Memory Organs," *U.S. Patent 2,815,488*, filed 4/28/54, issued 12/3/57, assigned to IBM.

[9] C.H. Bennett, "Logical Reversibility of Computation," *IBM J. Res. Dev.* vol. 17, 1973, pp. 525–532.

[10] J.A. Wheeler and W.H. Zurek, eds. *Quantum Theory and Measurement.* Princeton: Princeton Univ. Press, 1983. p. 782.

[11] Y. Lecerf, "Machines de Turing Réversibles. Récursive insolubilité en $n\epsilon N$ de l' équation $u = \theta^n u$, o ù θ est un 'isomorphisme de codes.' " *Compt. Rend.* vol. 257, 1963, pp. 2597–2600; "Récursive insolubilité de léquation générale de diagonalisation de deux monomorphismes de

monoïdes libres $\phi\chi = \psi\chi$," *Compt. Rend.* vol. 257, 1963, pp. 2940–2943.

[12] T. Schneider, "Use of Information Theory in Molecular Biology," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 102–110.

[13] H.S. Leff and A.F. Rex, "Maxwell's Demon: Entropy Historian," In: *Proc. Entropy Conference, April, 1994. Cambridge: MIT Press, to be published.*

[14] K.K. Likharev, "Classical and Quantum Limitations on Energy Consumption in Computation," *Int. J. Theor. Phys.* vol. 21, 1982, pp. 311–326.

[15] R.C. Merkle, "Towards Practical Reversible Logic," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 227–228; J.S. Hall, "An Electroid Switching Model for Reversible Computer Architectures," *ibid.*, pp. 237–247; J.G. Koller and W.C. Athas, "Adiabatic Switching, Low Energy Computing, and the Physics of Storing and Erasing Information," *ibid.*, pp. 267–270.

[16] C.L. Seitz, A.H. Frey, S. Mattisson, S.D. Rabin, D.A. Speck and J.L.A. van de Sneepscheut, "Hot-Clock nMOS," *1985 Chapel Hill Conference on VLSI.* Chapel Hill: Computer Science Press, pp. 1–17.

[17] A. Kramer, J.S. Denker, S.C. Avery, A.G. Dickinson and T.R. Wik, "Adiabatic Computing with the 2N-2N2D Logic Family," preprint.

[18] H.S. Leff and A.F. Rex, eds. *Maxwell's Demon.* Princeton: Princeton Univ. Press, 1990.

[19] R. Landauer, "Computation, Measurement, Communication and Energy Dissipation," In: *Selected Topics in Signal Processing.* S. Haykin, ed. Englewood Cliffs: Prentice–Hall, 1989, pp. 188-196. (Note: Some diagrams printed with wrong orientation.)

[20] D. Gabor, "Light and Information," In: *Progress in Optics.* E. Wolf, ed. Amsterdam: North–Holland Publishing Co. vol. I, 1961, pp. 109–153.

[21] F.T.S. Yu, *Optics and Information Theory.* New York: John Wiley, 1976, Ch. 5, pp. 93-111; A.F. Rex, "The Operation of Maxwell's Demon in a Low Entropy System," *Am. J. Phys.* vol. 55, 1987, pp. 359–362; J.D. Barrow and F.J. Tipler, *The Antropic Cosmological Principle.* Oxford: Clarendon Press, 1986, See pp. 179, 662; K.G. Denbigh and J.S. Denbigh, *Entropy in Relation to Incomplete Knowledge.* Cambridge: Cambridge Univ. Press, 1985, pp. 1–118; E.E. Daub, "Maxwell's Demon," *Stud. Hist. Phil. Sci.* vol. 1, 1970, pp. 213–227.

[22] R. Landauer, "Energy Requirements in Communication," *Appl. Phys. Lett.* vol. 51, 1987, pp. 2056–2058;

[23] W. Porod, "Comment on 'Energy Requirements in Communication ' " *Appl. Phys. Lett.* vol. 52, 1988, p. 2191; R. Landauer, "Response to 'Comment on Energy Requirements in Communication ' " *Appl. Phys. Lett.* vol. 52, 1988, pp. 2191–2192.

[24] C.M. Caves and P.D. Drummond, "Quantum Limits on Bosonic Communication rates," *Rev. Mod. Phys.* vol. 66, 1994, pp. 481–537

[25] R. Landauer, "Fundamental Limitations in the Computational Process," *Ber. Bunsenges. Phys. Chem.* vol. 80, 1976, pp. 1048–1059.

[26] P. Benioff, "The Computer as a Physical System: A Microscopic Quantum Mechanical Hamiltonian Model of Computers as Represented by Turing Machines," *J. Stat. Phys.* vol. 22, 1980, pp. 563–591; "Quantum Mechanical Hamiltonian Models of Turing Machines," *J. Stat. Phys.* vol. 29, 1982, pp. 515–546; "Quantum Mechanical Models of Turing Machines that Dissipate No Energy," *Phys. Rev. Lett.* vol. 48, 1982, pp. 1581–1585.

[27] R. Landauer, "Is Quantum Mechanics Useful?" *Proc. R. Soc. London A,* 1994, *in press*; "Is Quantum Mechanically Coherent Computation Useful?" *Proc. 4th Drexel Symposium on Quantum Nonintegrability, Sept. 8–10, 1994.* D.H. Feng and B.-L. Hu, eds. (International Press), *to be published.*

[28] R. Feynman, "Quantum Mechanical Computers," *Opt. News* vol. 11(2), 1985, pp. 11-20; [Reprinted in: *Found. Phys.* vol. 16, 1986, pp. 507–531.]

[29] Papers presented at the 1981 Conference on Physics of Computation, Part I: Physics of Computation, *Int. J. Theor. Phys.* vol. 21, 1982, pp. 165–350; Part II: Computational Models of Physics, *ibid.*, 1982, pp. 433–600; Part III: Physical Models of Computation, *ibid.*, 1982, pp. 905–1015.

[30] P. Benioff, "Quantum Mechanical Hamiltonian Models of Discrete Processes That Erase Their Own Histories: Application to Turing Machines," *Int. J. Theor. Phys.* vol. 21, 1982, pp. 177–201.

[31] J. Mehra, *The Beat of a Different Drum.* Oxford: Oxford Univ. Press, 1994, Ch. 24, pp. 525–542.

[32] R. Jozsa, "Computation and Quantum Superposition," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 192–194; A. Berthiaume and G. Brassard. "Oracle Quantum Computing," *ibid.*, pp. 195–199.

[33] L.B. Levitin, "Physical Information Theory Part I. Quasiclassical Systems," In: *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993, pp. 210–214.

[34] *Proc. Workshop on Physics and Computation.* Los Alamitos: IEEE Computer Society Press, 1993.