Notes on the history of reversible computation

by Charles H. Bennett

We review the history of the thermodynamics of information processing, beginning with the paradox of Maxwell's demon; continuing through the efforts of Szilard, Brillouin, and others to demonstrate a thermodynamic cost of information acquisition; the discovery by Landauer of the thermodynamic cost of information destruction; the development of the theory of and classical models for reversible computation; and ending with a brief survey of recent work on quantum reversible computation.

Concern with the thermodynamic limits of computation was preceded historically by the paradox of Maxwell's demon [1] and the realization that one bit of information is somehow equivalent to k ln 2 units of entropy, or about 2.3×10^{-24} cal/Kelvin. This equivalence was implicit in the work of Szilard [2] and became explicit in Shannon's use [3] of the term "entropy" and the formula

$$H = -\sum_i P_i \log P_i$$

to describe the self-information of a message source.

The history of this subject is noteworthy because it offers an example of how ideas that are strikingly successful in one area of science (in this case the uncertainty principle and the theory of black-body radiation) can stimulate unconscious false analogies, and so impede progress in other areas of science (thermodynamics of measurement and computation).

In the nineteenth century, despite the vision of Babbage, computation was thought of as a mental process, not a mechanical one. Accordingly, the thermodynamics of computation, if anyone had stopped to wonder about it, would probably have seemed no more urgent as a topic of scientific inquiry than, say, the thermodynamics of love. However, the need to think seriously about the thermodynamics of perceptual and mental processes was thrust upon science by the famous paradox of "Maxwell's demon," described as follows by its inventor, in a passage of admirable clarity and foresight [1]:

"One of the best established facts in thermodynamics is that it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or pressure without the expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up. But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have

REPRINTED FROM IBM JOURNAL OF RESEARCH AND DEVELOPMENT, VOL. 32, NO. 1, 1988; ©1988, 2000

0018-8646 / 00 / $5.00 © 2000 IBM

seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

"This is only one of the instances in which conclusions we have drawn from our experience of bodies consisting of an immense number of molecules may be found not to be applicable to the more delicate observations and experiments which we may suppose made by one who can perceive and handle the individual molecules which we deal with only in large masses.

"In dealing with masses of matter, while we do not perceive the individual molecules, we are compelled to adopt what I have described as the statistical method of calculation, and to abandon the strict dynamical method, in which we follow every motion by the calculus.

"It would be interesting to enquire how far those ideas about the nature and methods of science which have been derived from examples of scientific investigation in which the dynamical method is followed are applicable to our actual knowledge of concrete things, which, as we have seen, is of an essentially statistical nature, because no one has yet discovered any practical method of tracing the path of a molecule, or of identifying it at different times.

"I do not think, however, that the perfect identity which we observe between different portions of the same kind of matter can be explained on the statistical principle of the stability of averages of large numbers of quantities each of which may differ from the mean. For if of the molecules of some substance such as hydrogen, some were of sensibly greater mass than others, we have the means of producing a separation between molecules of different masses, and in this way we should be able to produce two kinds of hydrogen, one of which would be somewhat denser than the other. As this cannot be done, we must admit that the equality which we assert to exist between the molecules of hydrogen applies to each individual molecule, and not merely to the average of groups of millions of molecules."

Maxwell offered no definitive refutation of the demon, beyond saying that we lack its ability to see and handle individual molecules. In subsequent years Smoluchowski [4] partly solved the problem by pointing out that a simple automatic apparatus, such as a trap door, would be prevented by its own Brownian motion from functioning as an effective demon. He also remarked [5],

"As far as we know today, there is no automatic, permanently effective perpetual motion machine, in spite of molecular fluctuations, but such a device might, perhaps, function regularly if it were appropriately operated by intelligent beings. . . ."

This apparent ability of intelligent beings to violate the second law called into question the accepted belief that such beings obey the same laws as other material systems. Szilard, in his famous paper [2], "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings," attempted to escape from this predicament by arguing that the act of measurement, by which the demon determines the molecule's speed (or, in Szilard's version of the apparatus, determines which side of the partition it is on) is necessarily accompanied by an entropy increase sufficient to compensate the entropy decrease obtained later by exploiting the result of the measurement. Szilard was somewhat vague about the nature and location of this entropy increase, but a widely held interpretation of the situation, ever since his paper appeared, has been that measurement is an inevitably irreversible process, attended by an increase of entropy in the universe as a whole by at least $k \ln 2$ per bit of information acquired by the measurement. Later we shall see this is not quite correct: The measurement itself can be performed reversibly, but an unavoidable entropy increase, which prevents the demon from violating the second law, occurs when the demon erases the result of one measurement to make room for the next. The existence of an irreducible thermodynamic cost for information destruction (as opposed to information acquisition) was only clearly recognized three decades later by Landauer [6], and another two decades elapsed before Landauer's insight was applied to explain the demon without invoking any thermodynamic cost of measurement [7–9].

Ironically, Szilard came quite close to understanding the thermodynamic cost of information destruction. At the end of his paper, where he followed one version of his demon apparatus through a complete cycle of operation, he found that resetting the demon in preparation for the next measurement generated $k \ln 2$ of entropy. Unfortunately, he did not pursue this finding to the point of recognizing that information destruction is always thermodynamically costly, and that therefore no thermodynamic cost need be postulated for information acquisition.

Szilard's partial insight was lost as subsequent workers neglected resetting, and instead attempted to prove in detail the irreversibility of various measurement processes, particularly those in which the demon observes the molecule with light. The emphasis on measurement and neglect of resetting probably represented unconscious biases from everyday experience, where information is thought of as valuable or at worst neutral, and from quantum mechanics, which strikingly demonstrated the nontriviality of the

measurement process. The influence of quantum mechanics, particularly the quantum theory of black-body radiation, can be seen in a discussion of Maxwell's demon in Brillouin's influential 1956 book *Science and Information Theory* [10]:

"The essential question is . . . *Is it actually possible for the demon to see the individual atoms?* . . . The demon is in an enclosure at equilibrium at constant temperature, where the radiation must be black body radiation, and it is impossible to see anything in the interior of a black body. . . . The demon would see thermal radiation and its fluctuations, but he would never see the molecules.

"It is not surprising that Maxwell did not think of including radiation in the system in equilibrium at temperature T. Black body radiation was hardly known in 1871, and it was thirty years before the thermodynamics of radiation was clearly understood and Planck's theory was developed."

Brillouin goes on to consider a dissipative measurement scheme in which the demon observes the molecules by photons from a non-equilibrium source such as a hot lamp filament, concluding that to see the molecule, the demon must use at least one photon more energetic than the photons comprising the thermal background, thereby dissipating an energy of order kT in the process of measurement.

By the 1950s the development of the theory of computation by Turing and others had made it commonplace to think of computation as a mechanical process. Meanwhile the development of electronic digital computers had naturally raised the question of the ultimate thermodynamic cost of computation, especially since heat removal has always been a major engineering consideration in the design of computers, limiting the density with which active components can be packed.

The general folklore belief at this time, descended from Szilard's and Brillouin's analyses, is expressed in a remark [11] from a 1949 lecture by von Neumann, to the effect that a computer operating at temperature T must dissipate at least $kT \ln 2$ of energy "per elementary act of information, that is, per elementary decision of a two-way alternative and per elementary transmittal of one unit of information."

A major turning point in understanding the thermodynamics of computation took place when Landauer [6] attempted to prove this folklore belief and found he couldn't. He was able to prove a lower bound of order kT for some data operations, but not for others. Specifically, he showed that "logically irreversible" operations—those that throw away information about the previous logical state of the computer—necessarily generate in the surroundings an amount of entropy equal to the information thrown away. The essence of Landauer's argument was that such operations compress the phase space spanned by the computer's information-bearing degrees of freedom, and so, in order to occur spontaneously, they must allow a corresponding expansion, in other words, an entropy increase, in other degrees of freedom.

[This argument is not without its subtleties; for example, a many-to-one operation such as erasure may be thermodynamically reversible or not, depending on the data to which it is applied. When truly *random* data (e.g., a bit equally likely to be 0 or 1) is erased, the entropy increase of the surroundings is compensated by an entropy decrease of the data, so the operation as a whole is thermodynamically reversible. This is the case in resetting Maxwell's demon, where two equiprobable states of the demon's mind must be compressed onto one. By contrast, in computations, logically irreversible operations are usually applied to nonrandom data deterministically generated by the computation. When erasure is applied to such data, the entropy increase of the environment is not compensated by an entropy decrease of the data, and the operation is thermodynamically irreversible [7].]

About 1970, having read Landauer's paper and heard him talk, I began thinking about the thermodynamics of computation. Initially I assumed, as he, that at least some logically irreversible operations were necessary to nontrivial computation. However, as a side project, I experimented with simple computations that could be done without them. For example, I wrote a reversible program that used repeated subtraction to test whether one integer is divisible by another. Such experiments revealed a common pattern: The computation consisted of two halves, the second of which almost exactly undid the work of the first. The first half would generate the desired answer (e.g., divisible or not) as well as, typically, some other information (e.g., remainder and quotient). The second half would dispose of the extraneous information by reversing the process that generated it, but would keep the desired answer. This led me to realize [12] that any computation could be rendered into this reversible format by accumulating a history of all information that would normally be thrown away, then disposing of this history by the reverse of the process that created it. To prevent the reverse stage from destroying the desired output along with the undesired history, it suffices, before beginning the reverse stage, to copy the output on blank tape. No history is recorded during this copying operation, and none needs to be, since copying onto blank tape is already logically reversible; the reverse stage of computation then destroys only the original of the output, leaving the copy intact. My technique for performing an arbitrary computation reversibly is illustrated in **Table 1**, with underbars indicating the positions of the tape heads.

A proof of the thermodynamic reversibility of computation requires not only showing that logically irreversible operations can be avoided, but also showing that, once the computation has been rendered into the logically

reversible format, some actual hardware, or some physically reasonable theoretical model, can perform the resulting chain of logically reversible operations in a thermodynamically reversible fashion. Approaching the problem with a background of prior interests in biochemistry and computability theory, I saw an analogy between DNA and RNA and the tapes of a Turing machine. The notion of an informational macromolecule, undergoing transitions of its logical state by highly specific (e.g., enzyme-catalyzed) reversible chemical reactions, offered a felicitous model within which thermodynamic questions about information processing could be asked and rigorously answered. Within this theoretical framework it is easy to design an "enzymatic Turing machine" [7, 12] which would execute logically reversible computations with a dissipation per step proportional to the speed of computation. Near equilibrium, the machine would execute a slightly biased random walk, making backward steps nearly as often as forward ones. The backward steps would not result in errors, since they would be undone by subsequent forward steps. True errors—transitions to logically unrelated states—would also occur in any system with finite potential energy barriers, but their rate could be made small (in principle arbitrarily small) compared to the rate of logically correct forward and backward transitions. The enzymatic Turing machine is an example of a "Brownian" reversible computer, in which the non-information-bearing degrees of freedom are strongly coupled to, and exert a viscous drag on, the information-bearing ones, resulting in a dissipation per step proportional to the speed of computation.

Although there are no known general-purpose (i.e., universal) enzymatic Turing machines in nature, there are enzymes analogous to special-purpose Turing machines, notably RNA polymerase. This enzyme, whose function is to make an RNA transcript of the genetic information in one or more DNA genes, may be viewed as a special-purpose tape-copying Turing machine. Under physiological conditions the enzyme is driven hard forward, and dissipates about 20 kT per step; however, the operation of RNA polymerase is both logically and thermodynamically reversible, and it is routinely operated both forward and backward in the laboratory by varying the relative concentrations of reactants (nucleoside triphosphates) and product (pyrophosphate) [13, 14]. When operating backward the enzyme performs the logical inverse of copying: It removes bases one by one from the RNA strand, checking each one for complementarity with the DNA before removing it.

Edward Fredkin, at MIT, independently arrived at similar conclusions concerning reversible computation. Fredkin was motivated by a conviction that computers and physics should be more like each other. On one hand he was dissatisfied with a theoretical physics based on partial differential equations and continuous space-time. He felt it unreasonable to invoke an infinite number of bits of information to encode the state of one cubic centimeter of nature, and an infinite number of digital operations to exactly simulate one second of its evolution. By the same token he felt it wrong to base the theory of computation on irreversible primitives, not found in physics. To remedy this he found a reversible three-input three-output logic function, the "conservative logic gate" able to simulate all other logic operations, including the standard ones AND, OR, and NOT [15, 16]. He showed that conservative logic circuits can perform arbitrary computations by essentially the same programming trick I had used with reversible Turing machines: Do the computation, temporarily saving the extra information generated in the course of obtaining the desired answer, then dispose of this information by the reverse of the process by which it was created.

Table 1 Scheme for reversible computation.

Stage	*Contents of*		
	Work tape	*History tape*	*Output tape*
Forward	_INPUT	_	_
	WORK	HIST_	_
	_OUTPUT	HISTORY_	_
Copy output	_OUTPUT	HISTORY_	_
	OUTPUT	HISTORY_	OUT_
	_OUTPUT	HISTORY_	_OUTPUT
Reverse	_OUTPUT	HISTORY_	_OUTPUT
	WORK	HIST_	_OUTPUT
	_INPUT	_	_OUTPUT

Fredkin's displeasure with continuum models resembles Landauer's well-known displeasure [17] with mathematical operations that have no physical way of being performed, e.g., calculating the 10^{100}th digit of pi. These doubts, however, led Fredkin to pursue the radical goal of finding a fully discrete basis for physics, whereas in Landauer they merely inspired a certain aesthetic indifference toward nonconstructive mathematics.

Fredkin was joined by T. Toffoli (who in his doctoral thesis [18] had refuted, by counterexample, an accepted but erroneous proof that reversible cellular automata cannot be computationally universal), and later by Gerard Vichniac and Norman Margolus to form the Information Mechanics group at MIT. The activities of this group are largely responsible for stimulating the current interest in reversible cellular automata with direct physical significance, notably deterministic Ising models [19–21] and momentum-conserving lattice gases that support a macroscopic hydrodynamics [22].

A major step toward Fredkin's goal of finding a reversible physical basis for computation was his discovery of the billiard-ball model of computation [16]. This takes

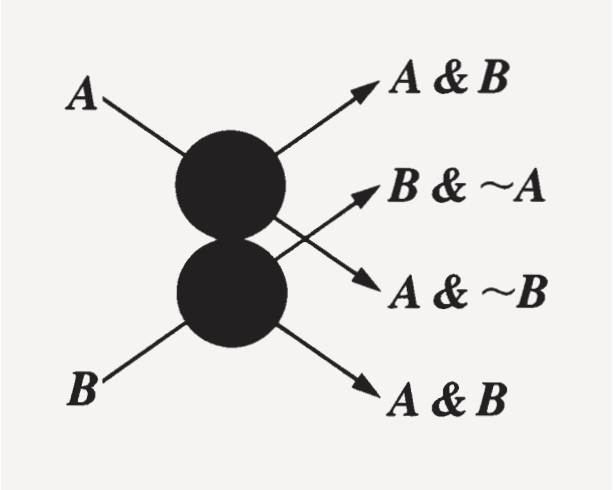

Figure 1

Use of a billiard-ball collision to realize a two-input, four-output logic function; data (1 or 0) represented by the presence or absence of a billiard ball on a given trajectory.

advantage of the fact that a collision between two classical hard spheres ("balls") diverts each one from the path it would have followed had the other been absent; thus a collision can be thought of as a two-input, four-output logic function whose outputs, for inputs A and B, are, respectively (cf. **Figure 1**),

A and B,
B and not A,
A and not B,
A and B.

Fredkin showed that, with the addition of "mirrors" to redirect the balls, such collisions can simulate any conservative logic function, and therefore any ordinary logic function. This implies that an infinite two-dimensional hard sphere gas, in an appropriate periodic potential (i.e., a periodic array of mirrors), is *computationally universal*—capable of being programmed through its initial condition to simulate any digital computation.

The billiard-ball computer is the prime example of a ballistic reversible computer. In contrast to the Brownian computers described earlier, ballistic computers operate with zero dissipation at finite speed, but they depend on isolating the information-bearing degrees of freedom from all sources of thermal noise, such as internal degrees of freedom of the balls or mirrors. Another way of characterizing the difference between Brownian and ballistic computers is to say that the former work by creating a low-potential energy labyrinth in configuration space, isomorphic to the desired computation, through which the system drifts despite thermal noise; the latter instead work by creating a dynamical trajectory isomorphic to the desired computation, which the system follows exactly in the absence of noise.

A number of other classical-mechanical models of reversible computation can be characterized as clocked Brownian models: The information-bearing degrees of freedom are locked to and driven by a master "clock" degree of freedom, with dissipation proportional to speed. These include the early coupled-potential-well models of Landauer and Keyes [6, 23], which were invented before the trick of reversible programming was known, but would function as Brownian reversible computers if reversibly programmed; the author's clockwork Turing machine [7], which invokes infinitely hard potentials to achieve zero error in a Brownian setting; Likharev's reversible computer based on Josephson junctions [24], which could probably be built, and Landauer's ball-and-pipe model [15, 25].

Returning to the question of Maxwell's demon, we can now give a detailed entropy accounting of the demon's cycle of operation. We refer to Szilard's [2] version of the demon, which uses a gas consisting of a single molecule. The demon first inserts a partition trapping the molecule on one side or the other, next performs a measurement to learn which side the molecule is on, then extracts $kT \ln 2$ of work by allowing the molecule to expand isothermally to fill the whole container again, and finally clears its mind in preparation for the next measurement. The discussion below of the classical Szilard engine follows [7]; an analogous quantum analysis has been given by Zurek [26].

According to our current understanding, each step of the cycle is thermodynamically reversible if we make the usual idealization that operations are carried out quasistatically. In particular, the measurement is reversible and does not increase the entropy of the universe. What the measurement does do, however, is to establish a correlation between the state of the demon's mind and the position of the molecule. This correlation means that after the measurement the entropy of the combined system (demon + molecule) is no longer equal to the sum of the entropies of its parts. Adopting a convenient origin for the entropy scale, the entropy of the molecule is one bit (since it may be, equiprobably, on either side of the partition), the entropy of the demon's mind is one bit (since it may think, equiprobably, that the molecule is on either side of the partition), but the entropy of the combined system is only one bit, because the system as a whole, owing to the correlation, has only two equiprobable states, not four.

The next phase of the cycle, the isothermal expansion, reduces the entropy of the environment by one bit while increasing the entropy of the demon + molecule system from one bit to two bits. Because the expansion destroys the correlation between demon and molecule (rendering the information obtained by the measurement obsolete), the entropy of the demon + molecule system is now equal to the sum of the entropies of its parts, one bit each.

The last phase of the cycle, resetting the demon's mind, reduces the entropy of the demon from one bit to zero, and accordingly, by Landauer's argument, must increase the entropy of the environment by one bit. This increase cancels the decrease brought about during the expansion phase, bringing the cycle to a close with no net entropy change of demon, molecule, or environment.

One may wonder how, in view of the arguments of Brillouin and others, the demon can make its measurement without dissipation. Though plausible, these arguments only demonstrated the dissipativeness of certain particular mechanisms of measurement, not of all measurements. In a sense, the existence of copying mechanisms such as RNA polymerase demonstrates the reversibility of measurement, if one is willing to call RNA synthesis a measurement of the DNA. More traditional reversible-measurement schemes can also be devised which are ideal in the sense of having no other effect than to establish the desired correlation between the measuring apparatus and the system being measured. Such a measurement begins with the measuring apparatus in a standard dynamical or thermodynamic state and ends with it in one of several states depending on the initial state of the system being measured, meanwhile having produced no change either in the environment or in the system being measured. **Figure 2**, for example, shows a classical billiard-ball mechanism based on the ideas of Fredkin that uses one billiard ball (dark) to test the presence of another (light) without disturbing the dynamical state of the latter. The apparatus consists of a number of fixed mirrors (dark rectangles) which reflect the billiard balls. First assume that the dark ball is absent. Then a light ball injected into the apparatus at X will follow the closed diamond-shaped trajectory $ABCDEFA$ forever, representing the value 1; conversely, the absence of the light ball (i.e., no balls in the apparatus at all) represents the value 0. The goal of the measurement is to inject another ball (dark color) into the apparatus in such a way that it tests whether the light ball is present without altering the light ball's state. By injecting the dark ball at Y at the appropriate time, the light ball (if present) is diverted from, but then returned to, its original path (following BGD instead of BCD), while the dark ball leaves the apparatus at M if the light ball was present and at N if it was absent.

One can design analogous mechanisms [7, 8] for reversibly measuring which side of Szilard's engine the molecule is on without otherwise disturbing the thermodynamic state of the engine or the environment. Such reversible nondemolition measurement schemes in general exist for classical systems, and for quantum systems in which the goal of the measurement is to distinguish among orthogonal states of the system, since these states may in principle be made eigenstates of an appropriate observable. Of course a quantum measurement cannot avoid disturbing a system which is presented to it in a superposition of eigenstates the measuring apparatus is designed to measure. The relation of irreversibility to quantum measurement has been considered by many authors (cf. the concise discussion in [27] and references therein).

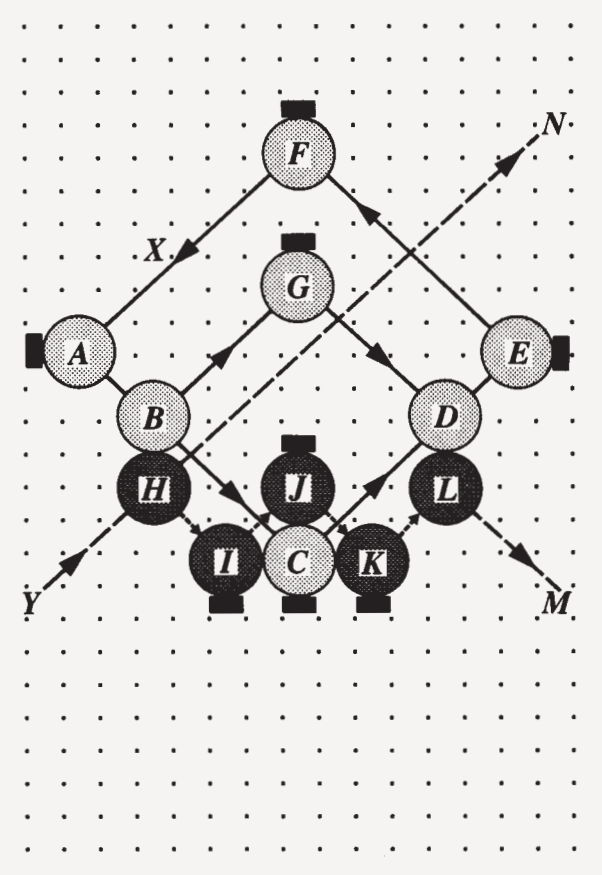

Figure 2

Reversible measurement in the billiard-ball model of computation.

An active research area recently has been the theory of quantum reversible computation. Chemical Brownian computers such as RNA polymerase are of course quantum systems, but because of the high temperature and short thermal de Broglie wavelength, quantum effects are subtle and quantitative (e.g., zero-point and tunneling corrections to reaction rates) rather than qualitative.

More distinctively quantum models have been considered by a number of authors [28–35]. These models are somewhat abstract by comparison with classical models, consisting typically of an array of two-state spins (each representing one bit) and a time evolution operator or Hamiltonian designed to make the spins pass through a sequence of states corresponding to the desired computation. The computationally relevant states are generally a subset of

a set of orthonormal "basis states," in which each spin is either up or down.

One of the earliest quantum models, by Benioff [28], used a Hamiltonian such that a basis state corresponding to the initial logical state of a reversible Turing machine would be transformed, at integer times, to basis states corresponding to successive logical successors. In other words, the Hamiltonian H was chosen so that the unitary operator U, representing evolution under the Hamiltonian for unit time, mapped each computationally relevant basis state onto its logical successor. Casting the logical time evolution of a Turing machine into the form of a unitary operator requires that all basis states have successors. Thus there can be no halt states, and all computations must be either infinite or cyclic. Since the Hamiltonian represents, albeit in an abstract form, the actual interactions among parts of the quantum computer that the designer is able to choose and control, Benioff considered it important for the Hamiltonian to be simple, and in particular not to depend explicitly on the global structure of the computation being performed. In order to achieve this, he found it necessary to make H time-dependent, in effect using a three-phase clock (two phases would also have sufficed) to turn on three Hamiltonians one after another, and making U the product of three non-commuting unitary operators $U = U_3U_2U_1$. In each of the clock phases, some of the spins (bits) in the computer flip conditionally on the state of others.

Feynman [29] found a way to define a simple, time-independent Hamiltonian for quantum computations: Instead of incorporating the direction of the computation (forward as opposed to backward) in the Hamiltonian, he incorporated it into the initial condition, which was now not a single basis state but rather a wave-packet-like superposition of basis states. The Feynman Hamiltonian for a given unitary transition operator U is of the form $H = U + U^+$, analogous to the Hamiltonian for a one-dimensional crystal in which spin waves can propagate either forward or backward according to their initial momentum. Feynman also noted that quantum computers can exhibit behavior intermediate between Brownian and ballistic: Thermal fluctuations in the Hamiltonian scatter the propagating computation wave like phonons in a crystal, so that under appropriate conditions the mean free path between scattering events is finite but much larger than one computation step. The computation then proceeds with net velocity proportional to the driving force, as in a Brownian computer, but with a proportionality constant that varies inversely with the mean free path, like electrical conductivity.

Zurek [30] compares the dynamical stability of quantum and classical ballistic computers with respect to errors in the initial condition ("software") and the Hamiltonian ("hardware"). In the billiard-ball model either type of error produces an exponentially growing error in the trajectory, whereas for quantum computers hardware-induced errors increase only quadratically with time and software errors do not increase at all.

Margolus [33] and Benioff [34] considered the problem of finding a universal quantum computer (with infinite memory) whose Feynman-type Hamiltonian nevertheless would have a finite range of interaction. For a serial computer such as a Turing machine, in which only one part is active at a time, this is not difficult; but when an analogous construction is attempted for a parallel machine such as a cellular automaton, Margolus found, on the one hand, that the finite range of interaction forbade synchronous updating of all the sites, and, on the other hand, that with asynchronous updating the computation no longer proceeded ballistically.

Deutsch [32] considered a more general kind of quantum computer that could be programmed to perform distinctively quantum operations such as generating two bits in an Einstein–Podolsky–Rosen superposition state. With Deutsch's computer it is possible to split a computation into two (or more) subtasks, perform the subtasks simultaneously in different Everett worlds, and then allow the results of the subtasks to interfere. An appropriate measurement on the final superposed state of the computer produces a probabilistic behavior of the output, which sometimes yields the desired answer (e.g., the exclusive-or, or some other linear function of the results of the two subtasks), and sometimes yields a "failure," an eigenstate of the output operator which says nothing about the results of the subtasks. Because of the probability of failure, quantum parallelization does not reduce the average time required to complete a parallelizable computation.

Landauer [35] has reviewed several quantum computation models in more detail than given here, pointing out some of the unphysical idealizations in existing models and the importance of specifying a quantum computer more concretely than by merely inventing a Hamiltonian.

Acknowledgments

I would like to acknowledge my early mentors Berni Alder and the late Aneesur Rahman, who had no direct interest in the theory of computation, but who taught me to think clearly about reversibility in statistical physics, resisting deception by common sense and discouragement by apparent paradoxes. I would also like to thank Rolf Landauer, under whose kind influence my interest in this subject took form nearly twenty years ago. Since then I have been inspired and assisted continually by many, especially Landauer, Gregory Chaitin, Edward Fredkin, Tom Toffoli, Paul Benioff, Norman Margolus, and David Deutsch.

References

1. J. C. Maxwell, *Theory of Heat*, 4th Ed., Longmans, Green & Co., London, 1875 (1st Ed. 1871), pp. 328–329.

2. L. Szilard, *Z. Phys.* **53,** 840–856 (1929).
3. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana-Champaign, IL, 1949.
4. M. von Smoluchowski, *Z. Phys.* (1912).
5. M. von Smoluchowski, lecture notes, Leipzig, 1914 (as quoted by Szilard [2]).
6. R. Landauer, *IBM J. Res. Develop.* **3,** 183–191 (1961).
7. C. H. Bennett, *Int. J. Theor. Phys.* **21,** 905–940 (1982).
8. C. H. Bennett, *Research Report RC-12526*, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598; edited version in *Sci. Amer.* (November 1987).
9. E. Lubkin, *Int. J. Theor. Phys.* **16,** 523–535 (1987).
10. L. Brillouin, *Science and Information Theory*, 2nd Ed., Academic Press, London, 1962.
11. J. von Neumann, *Theory of Self-Reproducing Automata*, Arthur Burks, Ed., University of Illinois Press, Urbana-Champaign, IL, 1966, p. 66.
12. C. H. Bennett, *IBM J. Res. Develop.* **17,** 525–532 (1973).
13. Judith Levin, Doctoral Dissertation, Biochemistry Department, University of California, Berkeley, CA, 1985.
14. George Kassaveties, *J. Biol. Chem.* **261,** 14256–14265 (1986).
15. R. Landauer, *Ber. Bunsenges.* **80,** 1048 (1976).
16. E. Fredkin and T. Toffoli, *Int. J. Theor. Phys.* **21,** 219–253 (1982).
17. R. L. Landauer, *IEEE Spectrum* **4,** 105 (1967).
18. T. Toffoli, *J. Comp. Syst. Sci.* **15,** 213–231 (1977).
19. G. Vichniac, *Physica* **10D,** 96 (1984).
20. Y. Pomeau, *J. Phys. A* **17,** 415 (1984).
21. M. Creutz, *Ann. Phys.* **167,** 62 (1986).
22. L. Kadanoff, *Physics Today* **39,** 7 (September 1986).
23. R. W. Keyes and R. Landauer, *IBM J. Res. Develop.* **14,** 152 (1970).
24. K. K. Likharev, *Int. J. Theor. Phys.* **21,** 311 (1982).
25. R. L. Landauer, *Int. J. Theor. Phys.* **21,** 283 (1982).
26. W. Zurek, *Frontiers of Nonequilibrium Statistical Physics*, G. T. Moore and M. O. Scully, Eds., Plenum Publishing Co., New York, 1986, pp. 15–161.
27. W. Zurek, *Ann. N.Y. Acad. Sci.* **480,** 89–97 (1986).
28. P. Benioff, *Phys. Rev. Lett.* **48,** 1581–1585 (1982); *J. Stat. Phys.* **29,** 515–545 (1982).
29. R. P. Feynman, *Opt. News* **11,** 11–20 (1985).
30. W. H. Zurek, *Phys. Rev. Lett.* **53,** 391–394 (1984).
31. A. Peres, *Phys. Rev.* **32A,** 3266–3276 (1985).
32. D. Deutsch, *Proc. Roy. Soc. Lond. A* **400,** 97–117 (1985).
33. N. Margolus, *Ann. N.Y. Acad. Sci.* **480,** 487–497 (1986).
34. P. Benioff, *Ann. N.Y. Acad. Sci.* **480,** 475–486 (1986).
35. R. Landauer, *Found. Phys.* **16,** 551–564 (1986).

Received April 8, 1987; accepted for publication October 5, 1987

Charles H. Bennett *IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598.* Dr. Bennett earned his Ph.D. from Harvard University in 1970 for molecular dynamics studies (computer simulation of molecular motion) under David Turnbull and Berni Alder. For the next two years he continued this research under the late Aneesur Rahman at Argonne Laboratories, Argonne, Illinois, coming to IBM Research in 1972. In 1973, building on the work of IBM Fellow Rolf Landauer, Dr. Bennett showed that general-purpose computation can be performed by a logically and thermodynamically reversible apparatus, one which is able to operate with arbitrarily little energy dissipation per step because it avoids throwing away information about past logical states. In 1982 he proposed a reinterpretation of Maxwell's demon, attributing its inability to break the second law to an irreducible thermodynamic cost of destroying, rather than acquiring, information. Aside from the thermodynamics of information processing, Dr. Bennett's research interests include the mathematical theory of randomness, probabilistic computation, and error-correction; the cryptographic applications of the uncertainty principle; and the characterization of the conditions under which statistical-mechanical systems evolve spontaneously toward states of high computational complexity. In 1983–85, as visiting professor of computer science at Boston University, he taught courses on cryptography and the physics of computation, and in 1986 was co-organizer of a conference on cellular automata held at MIT.