Preface

We live in a quantum world, in which probabilities, not certainties, govern what we see at the submicroscopic level. Interpretations of this fact have been the subject of endless debate since the formulation of quantum theory in the 1920s. One thing, however, is new: In the past decade, we have become increasingly familiar with quantum states. Indeed, at Los Alamos and other laboratories across the globe, individual quanta are being manipulated in ways only dreamt of before.

Those efforts have recently intensified as scientists are exploiting a newly identified aspect of the quantum world. It is called quantum information. Its smallest unit is the qubit, a two-level quantum system that can be measured to reveal a “yes” or “no” answer to a question. Thus, measurement of a qubit yields one classical bit of information. Under appropriate conditions, many systems behave as qubits. The polarization states of a single photon, spin-half nuclei in NMR experiments, and a system composed of two relatively stable levels of an ion are among the types of qubits explored at Los Alamos.

Unlike a classical bit, a qubit can be in a pure, yet superposed, state, in which it occupies both levels simultaneously. When measured, a superposition of the two levels behaves like a classical “probabilistic bit,” or pbit, yielding random yes or no answers according to the probability law associated with the particular measurement. The law has the following generic form: p is the probability of measuring yes, and $(1 - p)$, of measuring no. Neither the state of a pbit (that is, its probability law), nor the state of the qubit can be determined from a single measurement. Instead, an infinite sequence of measurements on independent but identically prepared copies of the system is necessary. However, when a qubit is prepared in a pure state, it has a property unknown in the classical world: There will always be one and only one independent property whose measurement will produce a yes answer with certainty, that is, with probability $p = 1$. Moreover, if that property is known, the probability laws associated with all possible measurements on the pure state are also known. In general, the pure states of a quantum system make up a complex projective Hilbert space, which for a qubit, can be pictured as points on the surface of a sphere. That is why qubits are often represented as vectors pointing along directions of a sphere.

$\langle ab\rangle - \langle ab'\rangle \le 1$
$i\hbar\dot{\psi} = H\psi$

In the illustration at right, a young man holds a qubit in his hands. His perspective rests on knowledge accumulated over the last century in quantum physics, information theory, and computer science, the fields that gave birth to the concept of quantum information. He symbolizes the potential of this new resource for communication and computation, as well as the curiosity and excitement it has generated among young men and women. Research on quantum information holds the promise of making quantum phenomena subject to control and manipulation of a new kind. It also holds the promise of bringing these phenomena into the classroom, where young people will grow up knowing the quantum first hand.

Inspiration is derived in many ways. At Los Alamos, a sense of history and the legacy of the great minds who were leading participants in the Manhattan Project are a continuing source. For that reason, this volume about the Los Alamos effort in quantum information and quantum science opens with thought-provoking words from John Wheeler and Richard Feynman (see pages vi–ix). Both were Manhattan Project pioneers, and as discussed below, both have helped launch the field of quantum information science and renew interest in the foundations of quantum theory and measurement.

$\lambda = h/p$
$E = \hbar\omega$
$i\hbar\dot{\psi} = H\psi$
$|\langle ab\rangle - \langle ac\rangle| \le 1 - \langle bc\rangle$
$S = -\mathrm{Tr}\,\rho\ln\rho$
$\Delta x\Delta p \ge \hbar/2$
$(\alpha|\uparrow\rangle + \beta|\downarrow\rangle)|0\rangle = \alpha|\uparrow\rangle|0\rangle + \beta|\downarrow\rangle|0\rangle$
$\mathrm{cnot} = |0\rangle_A {}^A\langle 0| + |1\rangle_A {}^A\langle 1|(|0\rangle_B {}^B\langle 1| + |1\rangle_B {}^B\langle 0|)$
$\rho = \Sigma_i p_i |\psi_i\rangle\langle\psi_i|$
$E = \hbar\omega$

The Strangeness of the Quantum World

"The elementary quantum phenomenon is the strangest thing in this strange world. It is strange because it has no localization in space and time. It is strange because it has a pure yes-no character—one bit of meaning. It is strange because it is more deeply dyed with an information-theoretic flavor than anything in all physics."
—John Archibald Wheeler (1984)

Wheeler is best known for working out the theory of nuclear fission with Niels Bohr in 1939 and for pioneering black-hole physics in the 1950s and 1960s. He has also spent well over half a century inspiring his many students and associates to think "outside the box." Together with Feynman, his graduate student in the early 1940s, Wheeler explored his "crazy" idea of treating particle trajectories going forward and backward in time on an equal footing. Both that experience and Dirac's ideas influenced the calculational shorthand known as Feynman diagrams and Feynman's formulation of quantum electrodynamics, for which Feynman received the Nobel Prize. In the 1960s and 1970s, Wheeler continually probed the connection between physics and information and opened the way for his graduate students and younger colleagues to help create a new field.

Quantum theory teaches us that, on the smallest scales, nature is observed to be granular. Electromagnetic radiation is absorbed and radiated in discrete units, which we call photons. The stable energy levels of an atom are also discrete, and electrons can be seen to go from one level to the next by "quantum jumps." The counterpoint to this ubiquitous discreteness is a form of continuity even more challenging to our everyday experience: Individual quantum systems can exist in a superposition of different states, corresponding, for example, to photons traveling along different paths in the famous double-slit experiment. Through measurement, the photon state, or wave function, "collapses" and becomes concentrated at the spot where it is observed, but repeating the measurement on another identically prepared photon typically produces a different, though equally definite, outcome. The state of each identically prepared photon is what determines the probability of obtaining different outcomes in many such repetitions and for many such measurements. To physicists imbued with the realistic local worldview of classical physics, the result is indeed surprising. Like a wide wavefront that has found its way through both slits simultaneously, each photon interferes with itself. Yet, when measured, as if by magic, each reduces to a point of light at some random location on the screen. The familiar interference pattern, predicted by both classical electromagnetic theory and quantum theory, arises when many photons are looked at together or in sequence (see the box "The Double-Slit Experiment" on page 142). The single photon—spread throughout space and observed only at a point in space—challenges our very concept of position as an attribute of the particle.

Viewed differently, in the delayed-

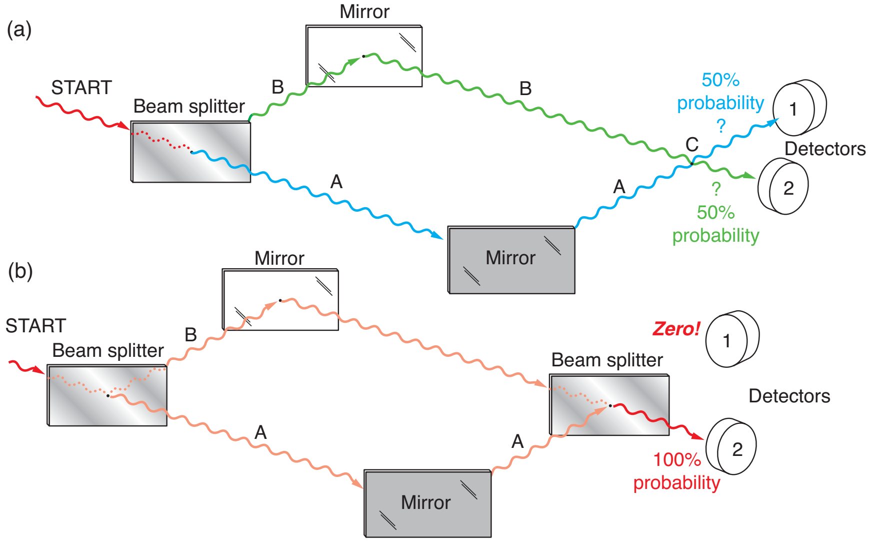


Figure 1. The Delayed-Choice Experiment
(a) At START, an incoming light wave encounters a beam splitter, which splits it into two beams of equal intensity. Each is reflected by a mirror and the two cross paths at point C. Detectors located past point C tell by which route an arriving photon has come. (b) The arrangement is the same as in (a) except that now a beam splitter is inserted at point C. It brings beams A and B into destructive interference on one side, so that detector 1 never registers anything, and into constructive interference on the other, so that every photon entering at START is registered at detector 2 in the idealized case of perfect mirrors and 100 percent photodetector efficiency. In (a), one finds out by which route the photon came. In (b), one has evidence that the arriving photon came by both routes. In the "delayed-choice" version of the experiment, one decides at the very last picosecond whether to insert the second beam splitter. In other words, one waits until the photon has done most of its travel before deciding whether the photon "shall have come by one route or by both routes." **(Diagram adapted with permission from John Wheeler, "Law without Law," in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek, Princeton, NJ: Princeton University Press, 1983.)**

choice experiment (Figure 1), the photon's behavior challenges our naive concept of causality. In (a), a photon hitting a beam splitter will follow path A or B, arriving at detector 1 or 2, respectively, with equal probability. One can deduce that, in this arrangement, the photon has followed a definite path: If either path is blocked, the count in the corresponding detector drops to zero. In (b), the setup is the same as in (a) except that a beam splitter is inserted at C, the point where the two paths cross. Now, interference causes all photons to arrive at detector 2 and none at detector 1. The photon's ability to traverse both paths is alone responsible for this situation: With either path blocked, the photons reach each detector equally. We can turn (a) into (b) by inserting a beam splitter at C, and we can choose whether to insert it at the very last moment. In this way, we can control whether the photon behaves as if it had taken one path or the other or had traveled along both paths. Now comes the contradiction to a local realist's view of causality: The beam splitter can be inserted after the photon is done traversing the region in question!

These paradoxes led Wheeler to view our physical reality through the lens of information theory: "Every item of the physical world has at bottom an immaterial source . . . what we call reality arises in the last analysis from posing yes-no questions and the registering of equipment-evoked responses; . . . in short, all things physical are information-theoretic in origin."

The link between what quantum mechanics tells us might happen—"multiple paths, interference patterns, spreading clouds of probability"—and what does indeed happen in the observable world is provided by the measurement process and/or the participation of the observer. The late Rolf Landauer of IBM, sometimes called the conscience of the physics of information, echoed Wheeler's view (1999): "I am suggesting that, contrary to our prevailing views, the laws of physics did not precede the universe and control it, but are part of it. Wheeler has stated that the laws of physics result from quantum measurement on the universe." These mind-bending thoughts about the elementary quantum phenomenon and the fundamental role of measurement, or information processing, in determining the laws of physics can be turned around to ask another profound question, "how does the nature of physical law limit our ability to process information?" Both questions were the subject of a seminal meeting in 1981.

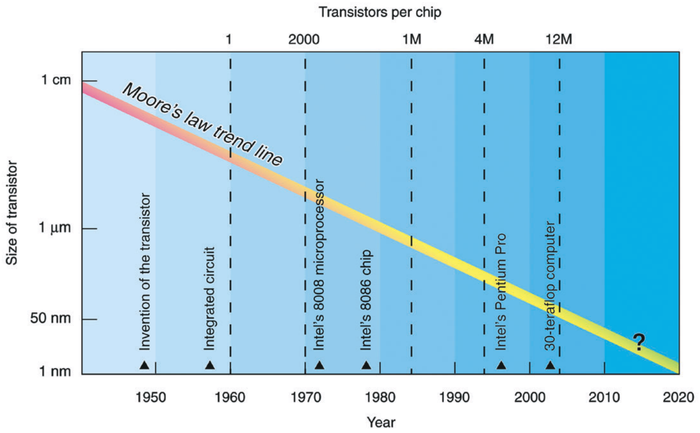


Figure 2. Miniaturization of the Transistor

By 1980, over 100,000 transistors were on a single chip. Today, that number is 40 million, or 20 million per centimeter squared. Quantum effects will become important when the size of the transistor approaches the nanometer scale and only a few electrons are involved in determining current flow.

Physical Limits on Computation and the First Models of a Quantum Computer

About 60 physicists and computer scientists gathered at the workshop "Physics of Computation" sponsored by the Massachusetts Institute of Technology (MIT). The organizers were Rolf Landauer, Tom Toffoli, and Ed Fredkin. The participants included Wheeler, Feynman, Charles Bennett, and Paul Benioff. According to the organizers, they all shared the belief that "physics and computation are interdependent at a fundamental level."

A general concern at that time was the drive toward ever-increasing computer power through miniaturization of components. Moore's law—the doubling of transistor density on a chip every eighteen months—had been describing an ongoing trend for several decades (see Figure 2). Because transistor elements were getting smaller, more and more of them were being crammed onto a chip, proportionately increasing computing power.

In the foreseeable future, each element would shrink to a size at which quantum effects become important. The question is how small could each get? Would the heat generated from so many computational steps in a tiny area lead to a literal meltdown of the chips? Could one use quantum mechanical elements to build computers—single atoms, perhaps? Would the time-energy uncertainty relation dictate the rate of energy dissipation? Would quantum fluctuations get in the way of reliability?

The research staff at IBM had

thought about these questions for many years and made some important strides. Landauer, for example, had repeatedly emphasized that information is always physical. He had delved deeply into the physics of information processing and in 1961 understood that erasure, the discarding of information, is an irreversible process that produces heat and increases entropy. He also assumed that computation necessarily involves erasure.

In 1973, Bennett showed that assumption to be false. Building on Szilard's work that connected information and entropy and Landauer's insight that erasure is the problem, he developed a logically reversible model of a Turing machine. This formal machine model of a universal computer had a memory tape, read-write head, and finite-state internal machine in the style of a Turing machine (see the box "The Universal Turing Machine" on page xvi). Bennett managed to design a reversible one-to-one mapping of information from input to output by employing three tapes instead of one: The first tape was for the input data; the second, for a history of intermediate results; and the third, for keeping a copy of the output. Because all operations would be done reversibly, the machine could run backwards, thereby retracing its steps, disposing of the intermediate results along the way, and returning to its initial state. This logical reversibility implied that, in principle, one could construct a thermodynamically reversible physical machine, which if run slowly, could perform any computation reversibly with arbitrarily little energy dissipation per step. Thereby, Bennett had found a way around the heat problem, but at the expense of speed. What, if any, were the limits quantum mechanics would place on computation?

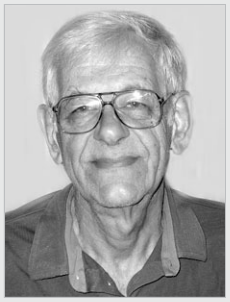
It was Benioff who first showed that reversible computation with no dissipation could be realized very naturally in a computer made of quantum mechanical parts. In 1980, he had begun developing quantum mechanical models of computation as a first step toward a model of intelligent systems. This very first model of a quantum computer consisted of a lattice of spin-half atoms that would evolve smoothly and deterministically according to the Schrödinger equation of quantum mechanics. Benioff invented an appropriate spin Hamiltonian that would govern the dynamics of this spin system, he proposed that the Hamiltonian act for a specific period to accomplish specified operations, and he showed that the states of the system would evolve with time, as needed to carry out the basic logic operations of a Turing machine. Because quantum mechanical time evolution is unitary, it generates a one-to-one reversible mapping of the system from one state to the next that can implement computational steps with no dissipation. Benioff presented his model at the 1981 "Physics of Computation" workshop. His ideas were revolutionary at the time. Many scientists had believed that any fast switching event would, by the time-energy form of the Heisenberg uncertainty principle, require a minimal energy expenditure, and therefore they expected to find intrinsic limits to the speed and accuracy in a computer obeying the laws of quantum mechanics.

"Information is inevitably tied to a physical representation and, therefore, to all the possibilities and restrictions allowed by our real physical universe. . . This is the viewpoint invoked by [Leo] Szilard. … His understanding of the physical nature of information was truly pioneering."
—Rolf Landauer (1999)

Benioff showed that this fear was unfounded: The laws of physics place no upper bound on the speed attainable or lower bound on energy dissipation during computation. The true significance of the uncertainty principle is, however, that the speed would be limited by the particular quantum dynamics of the computer: That is, the time per operation is limited by the Hamiltonian (energy) in the system divided by $\hbar$. Furthermore, the size of the elements could be reduced to individual atoms, or as we will see below, individual photons. Of course, although Hamiltonian evolution was simple to describe theoretically, Benioff had not dealt with the practical issues such as creating the initial state of the system, reading out the answer, the probabilistic nature of the quantum mechanical answer, and keeping the system isolated from the environment.

In "Zig-Zag Path to Understanding," Landauer recalls Benioff's 1981 presentation and the reaction to it: "[My own] attempts to produce a quantum version of the reversible Bennett-Fredkin-Turing machine had gotten hopelessly bogged down . . . Benioff saw the way to do that. You invoke a Hamiltonian (or a unitary time evolution) that 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. . . . Feynman was present at the 1981 workshop at MIT, where many of us discussed Benioff's notions. . . . Did we understand Benioff? Feynman did not need much of a clue, and as a result generated his own very appealing and

Quantum Issues at the 1981 Workshop as Remembered by Paul Benioff



"During the 1960s and 1970s, there was much interest in making fast, more powerful computers by miniaturizing components and packing more computer power into smaller volumes of space and time. However, there were two main problems: One was the appearance of quantum mechanical effects and the other was the generation of heat due to the irreversibility of the computation process. Until the work of Bennett in 1973, it was thought that the computation process was necessarily irreversible, with energy dissipation associated with information erasure. However Bennett showed that to every irreversible computation there exists an equivalent reversible computation."

"Yet Bennett's work did not address concerns related to quantum effects. Here, the concerns were twofold. One was that the energy-time uncertainty principle meant that the amount of energy dissipated per computation step was bounded below by Planck's constant divided by the switching time. However, as Landauer pointed out in 1982, the uncertainty principle does not mean that the energy is necessarily dissipated. I used this fact implicitly in my models. They operated at the quantum limit in that the total energy of the system was given by the energy-time uncertainty principle, but that energy was not dissipated."

"The other concern was that computation steps of a conditional nature—if a system is in state 0, do this; if it is in state 1, do that—necessarily involved measurement, which of course, does dissipate energy. The view that reading is equivalent to measurement is again erroneous. It ignores the fact that measurement consists of two stages: first establishing a correlation between states of the measured system and the apparatus, that is, an entangled state of the system and the measuring apparatus, and second, amplification or decoherence. It is this latter stage of decoherence, much studied and developed by Wojciech Zurek, that leads to dissipation. However, only the first step is necessary in quantum mechanical models of computation. This step, which does not dissipate energy, was used implicitly in my models and is an essential part of quantum computation models used today." (Private communication)

effective view of quantum mechanical computation" (Landauer 1994).

The keynote address at the 1981 workshop was delivered by Feynman, and it too was to have a profound impact on the community. The topic, tangential to the rest, was the problem of simulating physics with computers—in particular, simulating quantum physics. Feynman told his audience that this topic had a twofold interest: "learning something about the possibilities of computers, and also something about possibilities in physics." This interest was fueled by his close association with Fredkin, a proponent of the idea that space and time are discrete, not continuous, and that the Universe is, in essence, a giant digital computer.

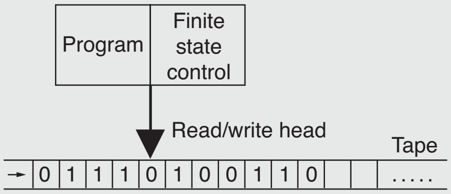
Feynman analyzed the problem with his typical flare and brilliance. He limited the computer to one with local interconnections and the type of simulation to one in which the number of computer elements required to simulate a large physical system is proportional to the space-time volume of the physical system. ". . . [C]lassical physics is local, causal, and reversible, and therefore apparently quite adaptable to computer simulation," provided, Feynman said, that we allow space-time to be discrete. In quantum mechanics, however, "we know immediately that we get only the ability, apparently, to predict probabilities . . ."

Could a system of probabilistic universal computers, classical Turing machines supplemented with random number generators, simulate the probabilistic world of quantum mechanics? His answer was a resounding "NO!" A probabilistic computer could not reproduce events with the same probabilities observed for quantum mechanical systems, without, in essence, simulating the entire universe at each point. A computer with only local interactions and polynomial resources would have to solve the famous hidden variable problem to match quantum probabilities, but John Bell had shown that only a nonlocal theory could predict the same probabilities as quantum mechanics (see the box "The EPR Paradox and Bell's Inequalities" on page xviii).

Feynman also concluded that such a classical computer could not simulate the wave function of a quantum system of *N* particles because the number of variables needed to describe the problem would grow faster than exponentially with *N*. He suggested, however, that "You can simulate it with quantum computer elements. It's not a Turing machine but

The Universal Turing Machine

In 1935, Alan Turing imagined a machine that would be capable of answering any question that could be answered logically. His invention was a blueprint for the modern programmable computer, and he proved theoretically that it could perform any computation that could be carried out through logical manipulations. The Turing machine has three elements (see figure): (1) an internal machine L that contains the program and can assume any one of a finite number of states, (2) a computation tape containing an infinite number of cells that serves as the memory, and (3) a read-write head that scans the tape, one cell at a time performs read/write operations on the cells, and can shift one cell to the left or right, or stay in place, depending on the contents of the cell, the state of the internal machine L, and the program instruction. The read/write alphabet is finite, say, zero and one, and it also includes a blank and a start symbol. Although operations such as addition require many steps, the machine is very powerful. The Church-Turing thesis states the following: The class of functions computable by a Turing machine corresponds exactly to the class of functions that we would naturally regard as being computable by any algorithm (definite procedure). Turing's invention was built on the insight of Kurt Gödel that both numbers and operations on numbers can be treated as symbols in a syntactic sense. Today, we take for granted that all information, including programmable instructions, can be expressed by strings of ones and zeros (or "yes" and "no" answers) and that all computations, from simple arithmetic to proving of abstract theorems, can be accomplished when a small set of mechanical operations (the program) are applied to these bit strings in some specified order.



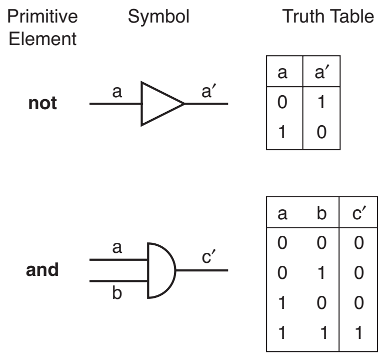
a machine of a different kind." Feynman then guessed that "every finite quantum mechanical system can be described exactly, imitated exactly, by supposing that we have another system such that at each point in space-time this system has only two possible base states. Either that point is occupied or unoccupied—those are the two states." In other words, a universal quantum simulator, closely resembling today's universal quantum computer, could be used to simulate discrete quantum systems. That idea is being pursued today. Only later, after Benioff's presentation at the 1981 workshop, did Feynman develop his own model of a universal quantum computer. It included a system for monitoring within the computer the progress of the calculation so that one would know the endpoint of the calculation and the time at which to read out the answer. At all decision points, however, this computer was in a definite state; never was superposition of different computational histories used as a tool.

In the spring of 1983, on the 40th anniversary of the Los Alamos National Laboratory, Feynman returned to Los Alamos for the first time since the 1940s. He joined his colleagues from the Manhattan Project era in a seminar on forward directions in physics. Feynman talked about reversible computing and his own model for a quantum computer in a talk entitled "Tiny Computers Obeying Quantum Mechanical Laws." Feynman's model was not the first, and it is not the model used in today's theoretical and experimental studies. Nevertheless, it stands as a record of Feynman's immense interest in this emerging area.

Physical Realizations of Classical vs Quantum Computers

One of the marvels of modern life is that the "universal" computer—any machine that is as powerful as a Turing machine—has become totally commonplace. All desktop computers are universal in the sense that, given enough time and memory, they can do any computation that is done on any other computer no matter how large. The model used to construct modern computers and the one used in this volume to describe quantum information processing are circuit models. For a classical computer, this model is realized as a set of physically distinct logic gates, usually implemented as transistor elements, connected by real wires. Information entered in the input register is processed as electric currents that go through a prescribed set of logic gates whose voltages are set according to the program instructions. Results are recorded in an output register. The figure at right shows two of the standard logic gates: the **not** gate and the **and** gate. Both the electronic symbols and corresponding truth tables for those operations in binary notation—0 = false (no), and 1 = true (yes)—are shown. These are the only gates needed to construct a universal computer that can perform all possible computations: In fact, a **nand** gate, constructed as an **and** gate followed by a **not** gate, suffices. Note, however, that the **and** (and **nand**) gate is obviously irreversible—one cannot determine the identity of the two inputs from the single output. A fully capable computer also needs fanout, the ability to send the same output to multiple inputs, and it needs to perform iteration (known as loops or recursion).

In the physical realization of reversible computing achieved with a quantum computer, on the other hand, there are no real wires. The input and output register is the same set of qubits, a row of, say, spin-half atoms in an ion trap, in a molecule, or embedded in a solid matrix. The "wires" carrying the qubits from one gate to the next are their time lines, and the logic gates are a sequence of unitary operators (typically external radio-frequency pulses and evolutions due to the internal interaction Hamiltonian of the system) that change the states of the qubits (see Figure 3). During the computation, the quantum mechanical wave function for the system evolves smoothly and deterministically according to the Schrödinger equation.



a	a′
0	1
1	0

a	b	c′
0	0	0
0	1	0
1	0	0
1	1	1

Once the computation is complete, the answer is obtained by a measurement, and, hence, is often probabilistic. A reliable answer typically requires repeated computations. It is, however, possible to design efficient quantum algorithms so that the final answer in the qubits is close to deterministic—any one measurement has sufficient information to allow extracting the desired answer with high probability. This deterministic feature is illustrated for the parity problem (introduced on page 21 of the primer) and Shor's algorithm (see the article "From Factoring to Phase Estimation" on page 38). The length of time for a quantum computation is limited by the intrinsic relaxation time of the system (various internal interactions can drive the two-level qubits to the ground state) and the decoherence time—the gradual "leakage" of quantum coherence to the environment.

Truth Table for **cnot** Gate

Q1	*Q2*	*Q1′*	*Q2′*
0	0	0	0
0	1	0	1
1	0	1	1
1	1	1	0

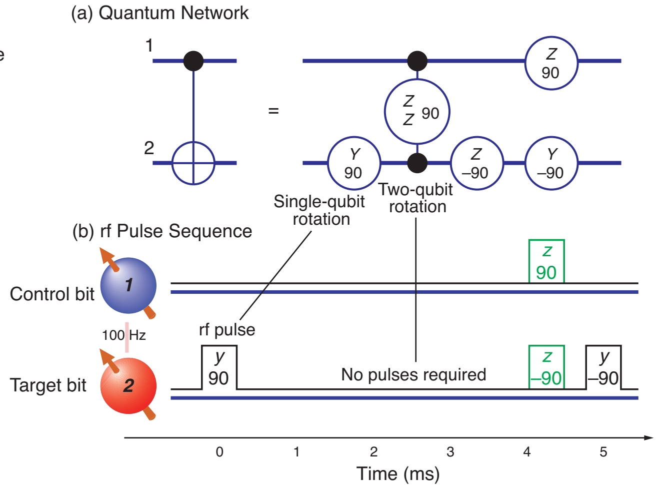


Figure 3. The cnot Gate in an NMR Quantum Processor
In quantum computers, the cnot (controlled not) gate plays a role analogous to that of the nand gate in classical computers. Part (a) shows the NMR quantum circuit, the sequence from left to right of 1-qubit rotations and 2-qubit internal Hamiltonian evolution, that executes the logic of the cnot gate, namely, reverse the spin of qubit 2 only if qubit 1 is in the 1 state (see Truth Table). Part (b) shows the rf pulse sequence needed to execute the cnot gate. Note that the two-qubit operation occurs by allowing the internal spin-spin Hamiltonian to evolve the system for a specified period. (See the article "NMR and Quantum Information Processing" on page 227.)

The Unique Properties of Quantum Information

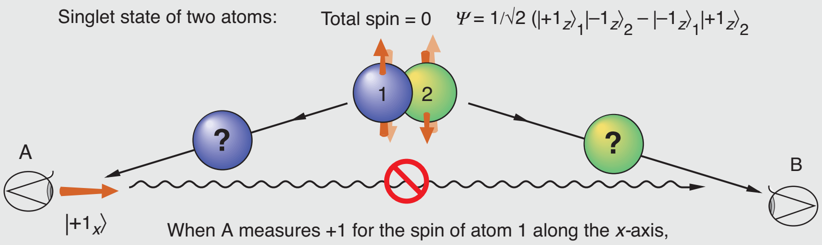
Feynman's notion that any finite quantum system could be simulated by a device made of spin-half atoms expanded the scope of what one might do with a computer made of quantum mechanical elements. In 1985, Deutsch took this idea one step further, suggesting that a computer made of elements obeying quantum mechanical laws could efficiently perform certain problem-solving and computational tasks for which no efficient classical solution was known. The key features of quantum mechanics to be exploited were the principle of linearity, which allows the components of a superposition of multiparticle states to evolve simultaneously, and the principle of interference, which allows certain superpositions

Continued on page xx

The EPR Paradox and Bell's Inequalities

When quantum theory was first formulated, Albert Einstein, Nathan Rosen, and Boris Podolsky questioned its completeness. They described a situation in which a predictable outcome could not be predicted by quantum mechanics. David Bohm illustrated this paradox (called EPR) using a molecule of two spin-half atoms with total angular momentum zero, that is, a coherent superposition of two product states (atom 1 in spin up) × (atom 2 in spin down) and (atom 2 in spin up) × (atom 1 in spin down)—see Figure A. For this maximally entangled state (defined in the main text), quantum mechanics predicts only that, if one atom's spin is measured along an axis chosen arbitrarily, the other atom's spin will always turn out to be its opposite when measured along the same axis. Quantum mechanics also requires, however, that each individual measurement have a random result. One concludes that, if the atoms are split apart and the spin of atom 1 is measured after the two are separated by a large distance, a measurement of the spin of the second atom along that same axis would be completely determined without any signalling from atom 1.

Figure A. Is Quantum Mechanics Complete?


Singlet state of two atoms:
Total spin = 0
$\Psi = 1/\sqrt{2}\ (|+1_z\rangle_1|-1_z\rangle_2 - |-1_z\rangle_1|+1_z\rangle_2$
1
2
?
?
A
B
$|+1_x\rangle$

When A measures +1 for the spin of atom 1 along the *x*-axis,
A can predict with certainty that B will measure –1
for the spin of atom 2 along that same axis even though A and B do not communicate.

In the worldview of a local realist, a complete theory is one in which every 'real' property of a system can be predicted. Further, if the outcome of a measurement can be predicted with certainty without interfering with the system, then the measurement result defines a 'real' property. Because the spin direction of atom 2 can be predicted with certainty, spin direction must be a 'real' property of the atom. Hence, if quantum theory were complete, it would predict the spin direction of each atom independently. In other words, "since the initial quantum mechanical wave function does not determine the result of an individual measurement, this predetermination (of the spin direction of atom 2) implies the possibility of a more complete specification of the state" (Bell 1964). To a local realist, therefore, the property of nonlocal correlation seen in David Bohm's example required introduction of a more complete theory, possibly involving "hidden variables" (or degrees of freedom over which one would have no control) that would determine the outcomes of individual measurements. This apparent incompleteness of quantum theory was one issue in the famous debate between Bohr and Einstein about the validity of quantum mechanics.

Any hope of a more complete theory was laid to rest when John Bell (1964) showed that no local realistic theory could possibly reproduce the probabilities computed according to quantum mechanics, without at some point invoking nonlocal effects. Figure B illustrates the basis of Bell's proof. In that figure, we construct a local realistic theory that matches the results depicted in Figure A. We imagine that the initial quantum state can be mimicked classically by two arrows pointing to random, but opposite, points on a sphere. When the system splits apart, each arrow has an equal probability of going to the right or left. With the simple local rules given in parts (i) and (ii) of Figure B, a classical theory can predict the perfect correlations seen when both spins are measured along the same axis—the spins point in opposite directions with probability 1. Only when one measures the two arrows (spins) in different directions, say, *z* and *d*, does the local realistic theory contradict the quantum results—see Figure B(iii). Clearly, classical analogues of entangled qubits need to be more complicated than a pair of arrows; as long as the results of measurements along all axes depend on a single variable, such as the arrow's direction in our example, the quantum mechanical results cannot be reproduced. Basically, qubits would need to be modeled by machines that calculate the results of measurements along different axes using different combinations of hidden variables, all of which may be random but must be correlated between the two measurement paths. This possibility was ruled out by Bell, who showed that the entire set of correlations implied in Figure B cannot be reproduced by any local realistic theory.

In every such classical system, one can ask for the probability $p(z+, d+, \underline{d}+)$ that the hidden variables of the first particle have such values that measuring its spin along the three fixed axes z, d, and $\underline{d}$ (see next page) would yield positive values. Because measurements along the $\underline{\mathrm{d}}$ axis can provide only two possible values, positive or negative—one concludes that

$p(z+, d+) =$
$\quad p(z+, d+, \underline{d}+) + p(z+, d+, \underline{d}-).$

On the other hand, one obviously has $p(z+, d+, \underline{d}+) \le p(d+, \underline{d}+)$ and $p(z+, d+, \underline{d}-) \le p(z+, \underline{d}-)$.

Putting these together, one obtains the master result that the distribution of hidden variables must satisfy

$$p(z+,\, d+) \leq p(z+,\, \underline{d}-) + p(d+,\, \underline{d}+) \quad .$$

Let us now consider an event in which the first spin is measured to be positive along the *z*-axis, and the second is observed to be negative along the *d*-axis. Because of the strict antiparallelism of the two spins whenever both are measured along *d*, we can conclude that, if the first spin had been measured along the *d*-axis, the measurement would have yielded a positive result. Thus, such events for spins 1 and 2 occur when, and only when, the hidden variable of the first spin is in such a state that it would provide positive results to measurements on both the *z*- and *d*-axis. In our notation, such a state for spin 1 happens with probability $p(z+, d+)$.

Thus, the probability of measuring a positive first spin along *z* and a negative second spin along *d*, $P_{zd}(+, -)$, is equal to $p(z+, d+)$. Using similar logic, we can transform our master inequality above to a statement about correlations:

$$P_{zd}(+,-) \leq P_{z\underline{d}}\,(+,+) + P_{d\underline{d}}(+,-) \quad ,$$

where $P_{zd}(+, -)$ represents the probability that, in an experiment in which the first spin was measured along *z* and the second along *d*, the observed outcomes were positive and negative respectively. This is a particular case of Bell's inequality, which every classical theory model must satisfy.

On the other hand, consider measuring the entangled system of two spin-half atoms along the same axes *z*, *d*, and $\underline{d}$. One can easily obtain the probabilities from quantum mechanics: $P_{zd}(+, -) = 3/8$, $P_{z\underline{d}}(+, +) = 1/8$, and $P_{d\underline{d}}(+, -) = 1/8$, and the inequality is clearly violated by the quantum system. Our classical reasoning led us astray: An entangled state is an indivisible unit, and trying to describe it probabilistically out of local properties assigned to its subsystems, even if they are correlated, is forever doomed to failure.

Bell's result changed forever our understanding of quantum mechanics and led to the modern view of quantum measurement.

Fixed Axes

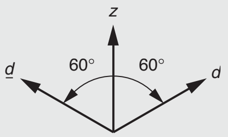
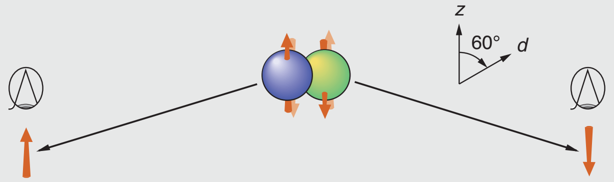


Figure B. Can a Local Realistic Theory Predict Quantum Mechanical Probabilities?

The following are assumptions for this thought experiment: (1) Spins in the initial state are assumed to point along opposite directions. (2) The spins fly off in opposite directions. (3) Each spin is equally likely to go to the left or right.

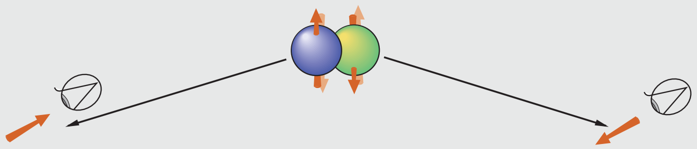
i. Measure both spins along the *z*-axis. To match quantum correlations, assume that the probability to measure a positive spin *p*(*z*+) or a negative spin *p*(*z*–) is 1/2. Both quantum and classical systems yield perfect correlations. The probability of spins pointing in opposite directions is 1.



Measurement Results

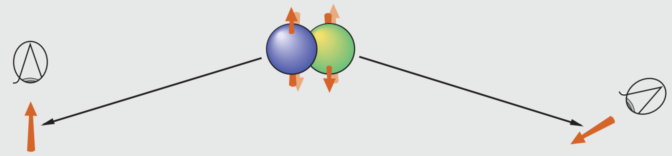
Correlation	Quantum	Classical
$P_{zz}(+,+)$	0	0
$P_{zz}(+,-)$	1/2	1/2
$P_{zz}(-,+)$	1/2	1/2
$P_{zz}(-,-)$	0	0

ii. Measure both spins along the *d*-axis. To match quantum correlations, add the deterministic rule that + along *z* is measured as + along *d* and that – along *z* is measured as – along *d*. Again, quantum and classical theory predict spins point in opposite directions with probability 1.



Correlation	Quantum	Classical
$P_{dd}(+,+)$	0	0
$P_{dd}(+,-)$	1/2	1/2
$P_{dd}(-,+)$	1/2	1/2
$P_{dd}(-,-)$	0	0

iii. Measure the spin traveling to the left along the *z*-axis and the one traveling to the right along the *d*-axis. Quantum and classical correlations do not match!



Correlation	Quantum	Classical
$P_{zd}(+,+)$	1/8	0
$P_{zd}(+,-)$	3/8	1/2
$P_{zd}(-,+)$	3/8	1/2
$P_{zd}(-,-)$	1/8	0

Continued from page xvii

to be detected with high efficiency.

To illustrate these ideas, first consider the properties of a single spin-half system, say, a single particle carrying the smallest nonzero amount of spin allowed by quantum mechanics. The particle behaves as a qubit: Measuring its spin (the amount of its intrinsic angular momentum) along any axis yields one of two quantized values, $+\hbar/2$ or $-\hbar/2$, and as mentioned early on, repeated measurements on independent, identically prepared systems yield a probability law for the two results (that is, the state of a pbit). The system, can, however, be prepared in a pure state with the spin pointing along a definite direction, so that measurement of the spin component along that direction results in $+\hbar/2$ with probability 1.

Suppose also that the system then evolves in isolation. By that we mean that laser pulses and other external sources can cause the spin to change direction in accordance with the laws of quantum mechanics, but because those sources have very large quantum uncertainties, the interaction with the qubit causes almost no change in their states. To put it differently, evolving in isolation means that the qubit can change state through interaction with the external world, but the external world has no information about the spin state of the qubit. Under these conditions, the pure state stays pure: Its spin always points in some definite direction (which, of course, changes with time), and if one happens to measure the spin along that direction or its opposite, one would be guaranteed to obtain a definite result and not disturb the state. Moreover, if one knew the preparation procedure and evolution that led to the state at the time of measurement, one could predict, in one stroke, both its direction and the probabilities for the measurement results along any direction.

For a register of N such qubits, the number of orthogonal pure states, or states in a complete basis, is exponentially large, 2^N to be precise. Just like a single qubit, however, this N-qubit quantum system can be in any superposition of these exponentially large number of basis states. Furthermore, the linearity of quantum mechanics implies that a sequence of few-qubit unitary operations designed to perform a given computation will do so on any superposition as easily as on a particular one. In this way, it effectively performs an exponentially large number of calculations simultaneously, without needing exponential resources at any stage.

When Deutsch conjectured (1985) that these simultaneous calculations could be exploited to solve problems more efficiently than could be done on a classical computer, he was quick to point out that this "quantum parallelism" is not an analogue of classical parallel computations. In fact, any such computation followed by measurement can yield only one N-bit answer. A direct measurement similar to that in case (a) of Figure 1 would collapse the final state to the results of a single randomly chosen calculation, a calculation that could have been performed on a classical computer equally easily. In contrast, quantum algorithms are carefully designed to be like case (b) of Figure 1; that is, interferences between the results of a large number of simultaneous evaluations are arranged so as to produce definite outcomes. Those outcomes provide information about global patterns (such as the periodicity of a function).

At first glance, this extra ability of quantum computers seems surprising. After all, if the initial superposition of the 2^N basis states is a collection of N pure qubits, each of which can be represented as an arrow pointing in some direction, and the computational steps maintain the purity of these individual qubits, then those steps could be viewed as rules for turning the arrows around. Such rules, it is easily shown, can be implemented on a classical computer with no difference in efficiency or precision. And if this were all there was to quantum computers, they could be no more powerful than classical ones.

Here, however, is the interesting part: Although quantum computations require only two-qubit operations at each step, many steps together are effectively multiqubit operations. Hence, the individual qubits do not evolve in isolation. Under these conditions, quantum mechanics assures us that only the entire N-qubit register is in a pure state, not the individual qubits; and this is where the miraculous nature of quantum correlations comes in. Many of the pure states of this N-qubit system display a peculiar phenomenon called entanglement: Even though the state of the register is pure—that is, we know as much as the uncertainty principle allows us to know about the system—and the entire system can be conveniently represented as a classical arrow, the states of the constituent qubits are not pure. And so, the state of the whole system is not describable by specifying the state of each qubit separately. An entangled state of more than one qubit is one that cannot be described as a probabilistic mixture of the product of single qubit states; a two-qubit state is called maximally entangled when it is pure, yet provides no information about local measurements on individual constituent qubits. An example of a maximally entangled state is provided in the box "The EPR Paradox and Bell's Inequalities" on page xviii. Entangled states are more akin to a classical register of probabilistic subsystems in which the interesting information (that is, the results of the calculation so far) is encoded in the numerous correlations between the subsystems. An analogous classical system, without the benefit of the multiparticle superpositions, would have to separately keep track of these correlations, which build up exponentially fast as the calculation proceeds. Whereas a quantum operation that changes the states of only a few qubits automatically updates the entire multiparticle superposition, the corresponding computational step in the classical system would require updating all these correlations and would become exponentially expensive. Note that it is not the entangled states per se that make quantum computation more efficient than classical analogues. Instead, enhanced computational power is a common feature of general quantum evolutions. Only a computation involving a very limited set of operations has the possibility to be mimicked classically. Conversely, unentangled evolutions of pure states can be mimicked classically because they, of necessity, involve very few kinds of operations. It is an open question whether the larger, but still limited, space of quantum evolutions that do not entangle mixed states of large number of qubits can be simulated efficiently classically, or whether they are powerful enough to perform scalable, useful computations.

Two specific features are responsible for the power of quantum computation: Because quantum mechanics causes multiparticle superpositions to evolve linearly, each computational step can carry out operations that would need an exponential number of classical resources. At the same time, the interference principle allows readout of certain global properties of the results. Those properties are often algorithmically unobtainable without evaluating the computation on each of the exponentially large number of input states. Deutsch's original quantum algorithm gave a solution for one such global property.

The area in which quantum entanglement does serve as a key resource is communication. The idea of exploiting the properties of quantum states for communication was born in the late sixties, when Stephen Weisner invented a quantum scheme for preventing counterfeiting of paper currency. His scheme was based on two properties of single quanta in pure states: First, though the results of measurements on quantum systems generally give random answers, a pure quantum system always provides a definite answer to some question. As a result, a quantum system is "unreadable" (in the sense of providing a definite result of measurement on it) to someone unaware of this question. Second, because a single quantum cannot be cloned (the no-cloning theorem), the system cannot be copied without having been read. Weisner's idea was to create serial numbers for paper currency by embedding in each bill a series of single-photon traps and filling them with a series of linearly polarized photons, each polarization standing for a particular number. If the series were composed of "nonorthogonal" (that is, prepared to answer different questions precisely) polarized photons, say, linearly polarized in both the horizontal/vertical directions and in the diagonal directions, then only the banks, which knew the precise directions to check, would be able to verify the number on the currency. Not having that specialized knowledge, counterfeiters would be unable to read or duplicate it without error. In fact, because measurement collapses the state to the observed result, any counterfeiter's attempt at reading the numbers could be detected by the bank with some probability.

Weisner's idea was ingenious though completely impractical. Yet, in the hands of Weisner's old college friend Bennett and Bennett's collaborator Gilles Brassard, it was transformed into a method for two parties to establish a secret encryption key while not allowing an eavesdropper to go undetected. One party creates a sequence of nonorthogonal photons, each polarized randomly either along the horizontal/vertical direction or along the diagonals, and sends them, one at a time, to the other party. The receiver can then measure each photon, randomly choosing one of the two bases. Because the sender can predict the measurement result only if the receiver and sender use the same basis, after the measurement the two need to communicate which basis each had used and discard the cases with different bases. Even if eavesdroppers listen to the conversation on a public channel and have access to the photon as it is being transferred, they can neither copy (clone) the photon (so as to store and measure it when its basis is finally announced) nor measure it in a random basis during transmission without affecting its polarization if they choose the wrong basis. The original parties always check the statistics of a small sample of the shared key to see if some process, or an eavesdropper, has affected the photons in flight and then use methods to insure, with high probability, the privacy of the shared key.

The central fact that single quantum systems in an unknown state cannot be cloned, or copied exactly, was proven by Bill Wootters and Wojciech Zurek, in 1982. Their elegant proof uses only the fact that quantum mechanics is a linear theory, in particular, that the principle of superposition always holds (see the box "The No-Cloning Theorem" on page 79). (Dennis Dieks proved the theorem independently that same year.)

Between 1985 and 1994, many people contributed to defining the specific elements of a universal quantum computer, to exploring categories of algorithms that might work more efficiently on a quantum computer, to developing applications of quantum information to communication, and in general, to developing the theory of quantum information in a way that paralleled the theory of classical information. But the interest was mainly confined to a relatively small group within the research community.

Then, without warning, the field broke wide open. Peter Shor demonstrated that finding the prime factors of an integer, a problem with great practical import, could be solved efficiently on a quantum computer. His solution took advantage of the mathematical fact that the remainders obtained when successive integral powers of any number *x* were divided by a fixed number *N* followed a cyclic pattern, and the corresponding period *r* was directly related to a factor of *N*. Shor's algorithm arranges an interference between the evaluations of a large sequence of these remainders so as to determine the period of the cycle with small error probability.

It is hard to overestimate how important Shor's work was for converting quantum computing and quantum information from an esoteric field involving only a few specialists to a field of general interest and real funding. One of the central problems in cryptography involves sending an encryption key when no private channel is available. Apart from the quantum key-distribution techniques described earlier, the best available methods in use today rely on the difficulty of factoring products of very large primes. To decrypt information, one has to find a solution to the so-called "discrete logarithm problem," whose practical solution calls for knowing the prime factors of an enormous number (see the box "Public-Key Cryptography: RSA" on page 72). Shor's proof that quantum computers could factor large numbers efficiently means that, if a quantum computer of sufficient power could be built, it would put at risk all such cryptographic methods. And these methods have been widely used to secure banking transactions, exchanges between intelligence agencies, and transactions over the Internet. Given the importance of his work, Shor was awarded the Nevalinna Prize for mathematical aspects of information science.

Both building a quantum computer and developing new cryptographic protocols such as quantum key distribution took on the aura of urgency. It seemed that these projects were not only interesting but also necessary from the point of view of security. Funding became available for mathematicians to find algorithms other than Shor's that could take advantage of quantum information. The most important one found to date is Grover's algorithm for unstructured searches. Many experimentalists were supported to try implementing what the mathematicians and theoretical physicists said could in principle be done. Ideas for constructing new qubits were cropping up everywhere. And excitement was generated in the popular press. But looming in the background was the certain knowledge that quantum states are fragile. Errors would inevitably occur, for example, through coupling to the environment. One had to find a way of preventing these without destroying the quantum states, which carry the information. That problem was solved in principle by Shor and Andrew Steane. They invented a scheme for error correction analogous to the strategies used for classical information. In 1998, Manny Knill, Raymond Laflamme, and Zurek proved the existence of an error bound, below which a quantum computation of arbitrary size could be implemented to arbitrary accuracy. Independent proofs of related results were done by Dorit Aharonov and Michael Ben-Or, Alexei Kitaev, and John Preskill. Implementing quantum computation in the laboratory became a realistic and compelling goal. Thus began a worldwide effort to build a quantum computer and to explore all the ways in which quantum information could impact science and technology. ■

Acknowledgments

The authors would like to thank Paul Benioff, Manny Knill, Salman Habib, and Wojciech Zurek for helpful comments and suggestions. One of the authors (Necia) would like to thank Tanmoy for sharing his deep understanding of quantum mechanics in this forum.

Further Reading

Bell, J. 1964. On the Einstein Podolsky Rosen Paradox. *Physics* **1** (3): 195.

Benioff, P. A. 1982. Quantum Mechanical Models of Turing Machines that Dissipate no Energy. *Phys. Rev. Lett*. **48** (23): 1581.

Bennett, C. H. 1973. Logical Reversibility of Computation. *IBM J. Res. Dev*. **6**: 525.

Deutsch, D. 1985. Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer. *Proc. R. Soc. London, Ser. A* **400**: 97.

Feynman, R. P. 1986. Quantum Mechanical Computers. *Found. Phys.* **16** (6): 507.

Landauer, R. 1994. Zig-Zag Path to Understanding. In *Proceedings of the Workshop on Physics and Computation PhysComp '94*. Los Alamitos, CA: IEEE Computer Society Press.

———. 1999. Information Is Inevitably Physical. In *Feynman and Computation: Exploring the Limits of Computers*. Edited by A. J. G. Hey. Cambridge: Perseus Books.

Lloyd, S. 1999. News and Views. *Nature* **400**: 720.

Wheeler, J. A. 1984. "Bits, Quanta, and Meaning." In *Theoretical Physics Meeting: Commemorative Volume on the Occasion of Eduardo Caianiello's Sixtieth Birthday*. Edited by A. Giovanni, M. Marinaro, F. Mancini, and A. Rimini. Naples, Italy: Edizioni Scientifici Italiani.

Additional references for John Wheeler and Richard Feynman are on pages vi–ix. For papers delivered at the 1981 MIT workshop Physics of Computation, see the *International Journal of Theoretical Physics* **21** (1), 1982.

Tanmoy Bhattacharya
Necia Grant Cooper
September 2002