

Time Travel and Computing

Hans Moravec

1991

The last few years have been good for time machines. Kip Thorne's renowned general relativity group at Caltech invented a new quantum gravitational approach to building a time gate, and, in an international collaboration, gave a plausible rebuttal of "grandfather paradox" arguments against time travel. Another respected group suggested time machines that exploit quantum mechanical time uncertainty. The technical requirements for these suggestions exceed our present capabilities, but each new approach seems less onerous than the last. There is hope yet that time travel will eventually become possible, even cheap.

Most time-machine fiction deals with the sociological implications of temporal trips or messages--indeed the time travel is often a mere literary device for placing humans in unusual situations. A recent paper from Thorne's group, by contrast, examines time travel at a basic physical level, deriving the quantum mechanical wave functions of systems that contain temporal loops. This article looks at the situation at an intermediate scale--the uses of time travel, packaged as negative time delay elements, on computation. This view is interesting because, on one hand, it predicts colossal improvements in the ability to solve important problems, and, on the other, provides a crisp logical metaphor for macroscopic implications of time travel.

A Brief History of Time Travel

When H.G. Wells wrote *The Time Machine*, his first novel, in 1895, the scientific world was unimpressed. In that Victorian age Science was in the process of crossing the t's and dotting the i's on Newtonian mechanics, and worrying about unemployment in the coming age of fully codified physical knowledge. Time was a rigid universal framework for the clockwork processes of physical law, without the slightest hint of a mechanism for evading its unvarying progress. That time machines were provably impossible was conveyed convincingly to generations of students, and echoes to this day.

The physics revolution at the beginning of the twentieth century shattered objective certainty about the immutability of time, but not most physicists' gut assurances. Special relativity combined space and time into a single continuum, with velocity a kind of rotation that transformed one into the other. A barrier--the speed of light--still separated the "spacelike" from the "timelike", but it was a fragile one. There was nothing in special relativity itself to preclude the existence of particles, now dubbed tachyons, that always moved faster than light. A tachyon message returned by a distant, rapidly receding relay could arrive before it was sent, a consequence construed by the conservative majority of the physics community as an indictment of tachyons. The impossibility of time travel (or closed causal loops, or, in the language of relativity, closed "timelike" loops) had become an axiom of physical law. Tachyons have not, in fact, been detected, even though they should be creatable with arbitrarily little energy (the faster they move, the less it takes), so perhaps the conservative majority is correct in this case (but perhaps they just have a tendency to hide - see below).

General relativity patches together tiny regions of flat special-relativistic spacetime into large gravity-warped structures. Powerful gravity fields imply radically convoluted spacetimes. Kurt Gödel was first, in 1949, to notice that general relativity predicted time travel under certain circumstances. In Gödel's solution to Einstein's

equations, the centrifugal tendency of a rotating universe exactly balances its tendency to gravitationally collapse. In such a universe the spacelike and timelike directions are skewed sufficiently that a spaceship accelerating around the universe can arrange to return to the place *and time* of its launch, giving the crew an opportunity to wave *bon voyage* to their departing younger selves. General relativity has been repeatedly confirmed experimentally in the large scale, so those who dislike the prediction take solace in the fact that our universe appears hardly to rotate.

The next major class of solutions, made in the 1960s by Roy Kerr and Ezra Newman and colleagues, are harder to dismiss. The Kerr-Newman solutions of the Einstein equations are for rapidly rotating and/or charged black holes. In the most extreme of these, the rotation of the body counteracts the gravitation enough to expose the twisted viscera of the black hole (normally hidden behind a discreet, one way, event horizon) including regions of *negative* spacetime, from which a spaceship could return to the outside universe before it entered. Attempts by the conservative majority to find independent reasons for a *cosmic censorship* rule to prevent such lewd exposure have been unsuccessful, so far. Their only comfort is that it would take about the mass of a galaxy, with extraordinary spin, to make a practical time machine this way.

In 1974 Frank Tipler published another solution to the general relativity equations, this time for the region around an extremely dense, very long rapidly spinning cylinder--dense as a neutron star, the diameter of a city block, with its surface moving at about one fourth the speed of light, and infinitely long, because that simplified the mathematics. Spacetime wraps itself around such an object like a roll of paper, producing alternate layers of negative and positive spacetime. A carefully aimed spaceship could swing around such an object, staying mostly in a negative region, and come out before it left. A finite length cylinder should also work, and might allow a time machine with only the mass of a star. But, the conservatives say, maybe it's not possible to prevent a finite cylinder from gravitationally collapsing lengthwise.

As yet, no one has devised a satisfactory comprehensive single theory that combines gravity and quantum mechanics--many try, and the implications of such a theory promise to be awesome. In 1988 Kip Thorne and company, patching together partial theories, described a time machine using both quantum mechanics and general relativity. A tiny, spontaneously formed, gravitational spacetime wormhole is pulled out of the hyperactive froth that is the quantum vacuum, and stabilized, by two large conductive plates, resembling an electrical capacitor. Initially these plates are as closely spaced as possible, and each becomes host to one "mouth" of the wormhole. When they are separated in our normal spacetime, they yet remain connected through the wormhole, which is an independent spacetime tunnel. Regardless of their external separation, a message or object entering one mouth appears instantly (by its own reckoning) at the other, as if the mouths were the two sides of a single door. Thorne's group then uses special relativity to differentially age the two mouths. One is taken on a "twin paradox" round trip at near the speed of light, so that less time elapses for it than its stationary counterpart. When it returns, the external separation between the two mouths has a time as well as a space component. A message sent into the itinerant mouth exits from the stationary one after a delay. And a message delivered into the stationary mouth exits from the traveller before it was sent! This kind of machine could perhaps be constructed with a planet's worth of aluminum spread out into plates the area of Earth's orbit, separated by the size of an atom--still beyond our means, but getting closer. The non-linear equations of general relativity are notoriously hard to solve, and only the very simplest cases have been explored. Even more significantly, there isn't any theory of quantum gravity yet at all. It's quite possible that in all this unexplored territory, waiting to be discovered, lie quite feasible ways to build time machines.

Another approach to time travel asks why it isn't observed routinely. There is no intrinsic time direction in Newton's mechanics nor in the differential equations of the new physics. The future determines the past just as fully as the past determines the future. Why, then, can our past selves leave messages for our future, but seemingly absolutely never the other way around? This question has not been definitively answered. Attempted explanations involve "boundary conditions", the initial values of physical quantities at the edges of space and time, which the differential equations of physical law then fill out. The universe must be somehow very different at one end than the other, and this difference orients the arrow of time. Most common is the thermodynamic explanation, which talks about the state of disorder or "entropy" of matter and energy. The universe started in a very rare, highly ordered state, and is running down into increasingly common states of disorder. Though this

explains why a ship can't run its engines by separating water into steam and ice, it does not explain why one can't send today's lottery numbers into yesterday by expending a few megawatt hours of energy.

One explanation which *does* was offered by John Wheeler and Richard Feynman in 1945. They noted that Maxwell's equations, the first "modern" physical theory, give two solutions for the effect of accelerating an electric charge. One, called the retarded wave, follows the acceleration and describes an electromagnetic disturbance diverging outward at the speed of light--the radio waves which link our civilization. The other, called the advanced wave, is for a similar disturbance that precedes the acceleration and converges on it (or, in another way of looking at it, diverges backwards in time). This latter wave is never seen. Wheeler and Feynman's analysis assumes that the advanced wave is, in fact, produced, and expands outwards into the past. There, eventually, it encounters a condition, perhaps the extreme density of the big bang at the beginning of the universe, that reflects it, producing a retarded wave exactly out of phase, that retraces its spacetime and exactly cancels it out. The retarded wave from an accelerated charge is not cancelled in an analogous manner, because there is no reflector in the future of the universe to reverse it--perhaps because the universe is "open", and expands forever. If this is a correct explanation, it might be possible to send signals backward in time by means of a kind a reflector that acts like the big bang--maybe a black hole.

Such a reflector would reverse and return a retarded wave, cancelling and thus apparently preventing it. If the reflector were installed one light year away from a light source aimed at it, it would suppress light from the source one year *before* installation. Similarly the suppression would go away one year before the reflector was removed. Messages could be sent a year into the past simply by moving the reflector in and out of the beam. Recently John Cramer devised an interpretation of Quantum Mechanics, called the transactional model, that uses this approach to explain every interaction. Transmission of a photon from one place to another actually involves two signals, one moving forward in time, the other backward, a "handshake", between the two locations. In Cramer's model a time communicator could be built that works much like the Wheeler-Feynman type, but using a wave absorber to prevent a transaction, rather than a reflector to cancel it.

There is a spookier possibility. Suppose it is easy to send messages to the past, but that forward causality also holds (i.e. past events determine the future). In one way of reasoning about it, a message sent to the past will "alter" the entire history following its receipt, including the event that sent it, and thus the message itself. Thus altered, the message will change the past in a different way, and so on, until some "equilibrium" is reached--the simplest being the situation where no message at all is sent. Time travel may thus act to erase itself (an idea Larry Niven fans will recognize as "Niven's Law"). This situation can be modeled quantum mechanically. If the message is a particle sent to the past, the wave function for that particle will subsequently propagate into the future, where it encounters and "interferes" with its original self, cancelling or reinforcing depending on the relative phase (which depends on the round trip length and other things). Now, the resultant wave function indicates the possible places the particle might actually be found in a measurement--large magnitudes are likely locations, near zero values unlikely ones. A round trip that causes the particle wave function to be cancel itself means that a particle is unlikely ever to be positioned to start the trip. It is exactly this kind of quantum probability effect (without reference to time travel) that powerfully confines electrons to discrete shells about the atomic nucleus and causes the "Pauli exclusion principle" that prevents two indistinguishable electrons from being in the same place simultaneously.

The Wheeler-Feynman advanced wave theory, Cramer's transactional model of quantum mechanics, and even quantum electrodynamics, the most accurately verified physical theory we have, all involve interactions that reach backwards as well as forwards in time. It may well be that time travel is as common as dirt, and shapes our physical laws, but conspires, by wave function interference, to prevent any operations that would result in logically inconsistent situations, and so makes overt macroscopic time travel difficult. Even with a time machine you will never succeed in preventing your own birth, or changing the antecedents of any present observation--some odd co-incidence or accident (perhaps one disabling your time machine) will always thwart the attempt.

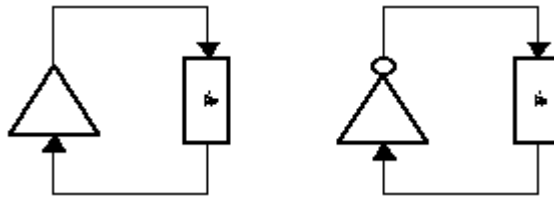
But this does not rule out carefully contrived *logically consistent* causal loops (The recent paper by Friedman, Morris, Novikov, Escheverria, Klinkhammer, Thorne and Yurtsever examines the consequences of such constraints on the wormhole time machine mentioned earlier). Wave interference typically appears as banded

patterns with a strong central (zero order) peak surrounded by a dark fringe, itself surrounded by the another (first order) bright fringe, and so on. If causal loops behave this way, it will be necessary to make major perturbations in experimental arrangement to skip from the usual zero order (no overt time travel) situation to a first order case of a non trivial consistent causal loop.

Let's suppose that eventually some approach results in a practical, even cheap, way to package time machines to make *negative time delay elements* --which output signals that predict what their inputs will receive some fixed time later. Such devices would have very interesting consequences for computational problems, which increasingly means almost every field of activity.

Time Loop Logic

Computer circuitry is composed of devices called gates that combine binary signals to produce other such signals. The simplest gate is an amplifier whose output is identical to its input. Almost as simple is the NOT gate, or inverter, whose output is 0 when its input is 1, and vice versa. All gates take a little time to respond to changes in their input--typically a few billionths of a second. When the input of an amplifier is connected to its output, the circuit will lock up with its output permanently at either 0 or 1. A NOT gate in a similar loop tends to oscillate rapidly between 0 and 1, at a rate that depends on its delay. It is possible to slow down this oscillation by inserting extra delay into the loop. Conversely, a negative delay element would speed up the oscillation. Imagine the two cases of an amplifier and a NOT gate with a certain delays each with its output connected to its input through a negative delay device:



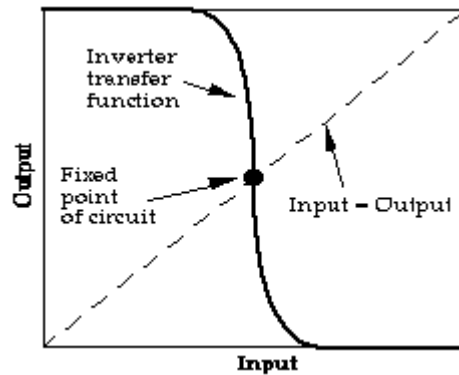
Causal loops involving an amplifier (triangle) and inverter (triangle with circle) in causal loops with negative time delays ($-t$) that cancel their forward delay.

The amplifier circuit is in a consistent causal loop--when first switched on, it can permanently assume either 0 or 1 without contradiction. The loop with the inverter, on the other hand, is a simple case of the classical time travel grandfather paradox, a paradoxical causal loop. An input of 1 to the inverter gives an output of 0, which is brought back in time to contradict the input. It takes some quantum mechanics to make sense of the situation, and we will have to say something about how the signals are physically represented. Most digital circuits represent signals as electrical voltages or currents in wires, which is inconvenient because electrons interact with each other and with matter in complex, hard to analyze ways. Some experimental circuits use much simpler space-crossing beams of light. Let's suppose 1 and 0 are encoded as coherent light beams of opposite phase (perfectly out of step with one another--one crests where the other has troughs). In that case a 0 that meets a 1, as in the inverter circuit, will simply cancel. Either alternative would have zero net probability, and the circuit (perhaps containing a charged laser, ready to emit a beam) should simply fail to turn on (ignite) at all, somewhat like a ball balanced on a knife edge that, against all odds, teeters indefinitely instead of falling to one side or the other. This is Niven's law at work in the small. The circuit finds itself perpetually in a dark fringe of an interference pattern.

On the other hand, suppose the two signals are encoded as beams of identical phase but opposite polarization (horizontal or vertical direction of vibration). Then the light in the inverter circuit will be in step with itself, and so should ignite. But what about the polarization? Neither of the possibilities individually results in a consistent situation, but quantum mechanics allows these to overlap in a *superposition*. The superposition that contains equal measures of 0 and 1 is changed by the inverter to an indistinguishable superposition of 1 and 0, and so

allows the circuit to remain consistent. Light with equal amounts of vertical and horizontal polarization is *unpolarized*. The inverting causal loop insures that the light remains perfectly unpolarized. An attempt to change this, for instance by slipping a polarizing filter somewhere into the circuit, should cause the beam to extinguish, as for phase modulation.

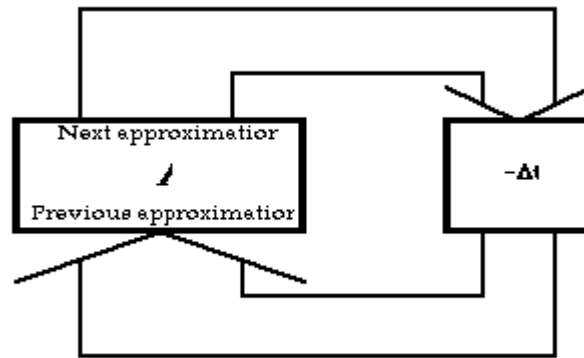
Imagine that the beam in the polarization modulated inverter circuit is sampled, perhaps by a partially silvered mirror, and examined. Quantum mechanics tells us that each extracted photon will be in a superposition of polarization until the polarization is measured, at which point it "collapses" to one possibility or the other. Sometimes the photon is in a 1 state, equally often it is found to be a 0. But suppose, as in existing computer circuitry, signals are represented not by single particles but by bulk aggregates which behave almost like classical continuous variables. An aggregate of random values will act like a signal intermediate in value between 0 and 1, say around 1/2. In classical terms, it makes sense that a circuit that converts 0 to 1 and 1 to 0 should leave some intermediate value unchanged:



There are other ways to view the situation. An inverting loop without a negative delay will oscillate between 0 and 1 at a frequency that depends on the time delay of the inverter. As we insert increasingly long negative delays in the circuit, the frequency of the oscillation increases. The superposition is when the frequency reaches infinity, and the circuit is in the 0 and 1 states simultaneously. Note also that the inverter's 1/2 state is extremely unlikely--a small deviation in the input would cause the output to saturate at one of the two extremes. Yet the circuit indefinitely maintains precisely that unlikely situation. The negative delay connects regions of spacetime in the same way that the orbit of an electron around an atomic nucleus connects different parts of its wave function, and so creates electron shells separated by empty regions. Like a focussing lens that brings to a single spot photons that would otherwise have landed all over a target, it produces interference patterns that make a few states of the world very likely and the rest unlikely. It's just that, by eliminating the standard possibilities, negative time delays make likely things that are otherwise nearly impossible.

Infinite Iteration

Here's a more complicated case. Make a computing box that accepts an input, which represents an approximate solution to some problem, and produces an output that is an improved approximation. Conventionally you would apply such a computation repeatedly a finite number of times, and then settle for the better, but still approximate, result. Given an appropriate negative delay something else is possible:



In this arrangement the result of each iteration of the function is brought back in time to serve as the "first" approximation. As soon as the machine is activated, a so-called "fixed-point" of F , an input which produces an identical output, usually signaling a perfect answer, appears (by an extraordinary coincidence!) immediately and steadily, just as either 1 or 0 appears in the simple amplifier loop. If the iteration does not converge, that is, if F has no fixed point, the computer outputs and inputs will shut down or hover in an unlikely intermediate state, like the inverting loop.

The Compleat NP Machine

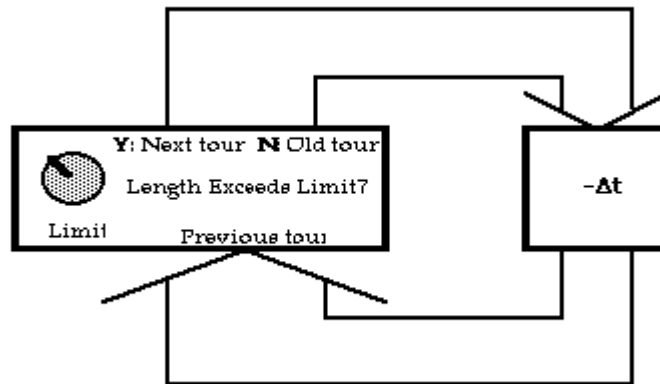
Generally speaking, whether a problem can be solved in a computer depends on its *computational complexity*, which describes how the difficulty grows as the problem increases in size. Finding the largest number in a list of numbers takes time proportional to the list length. A graph of computing time versus list length yields a straight line--the complexity is *linear*, and this is an easy problem. Sorting the list into numerical order is harder. Simple methods take time proportional to the square of the list length--the graph is an upward curving parabola--and even the best possible sorting methods give slight upward curves. The cost of solving a system of n equations in n unknowns grows as the cube of n , and other problems grow as the fourth, fifth or even higher powers of the problem size. But any problem whose difficulty can be expressed as a fixed power of size is of *polynomial* complexity, and even large instances are soon solvable in a world where computer power doubles every few years.

Not so problems of *exponential* complexity. These multiply in cost with each fixed increment of problem size. Whereas a linear problem grows as 1, 2, 3, 4, 5, 6, ..., and a cubic as 1, 8, 27, 64, 125, 216, ..., an exponential problem's cost may grow as 1, 10, 100, 1000, 10000, 100000 and so on to the astronomical. An important class of apparently exponential problems are called NP, short for Nondeterministic Polynomial, meaning exponential problems that could be solved in polynomial time given an exponential supply of computers to examine alternatives. Examples include many design problems such as finding the best arrangement of logic gates, or the fastest program, to compute a given function, and also finding limited length proofs of mathematical theorems. Solving such problems could synergistically increase the power of the machines that solve them. Even poetry or music writing might fit, if approached as the problem of finding the best sequence of constrained words or sounds to express an idea or a state of mind.

The hard core of NP problems are called NP complete, and it's been shown that a fast solution for any NP complete problem can be translated into a fast solution for any other. One famous and convenient NP task is the *traveling salesman problem* --given n cities and the distances between pairs of them, find the shortest tour that passes through each city exactly once.

A version of the time loop iteration box of the last section can quickly solve such a task. The F box for this problem takes as input a particular tour, that is, a permutation of the cities. It also has a knob whose position

specifies a limit on the length of the tour. The box calculates the input tour's length, and outputs the same tour if that length is less than or equal to the limit. If the length exceeds the limit, the box generates the next permutation (by some permutation counting scheme). When the circuit is activated it settles down immediately to a tour shorter than the limit, or sticks in an undecided state if no such solution exists--i.e. if the limit was set too low.



Turning the knob to find the boundary between these two conditions solves the problem. The "knob turning" can, of course, be done automatically by additional machinery, for instance an ordinary computer. In fact, one can imagine building a very general "chronocomputer" by simply slipping a sufficiently large "negative delay register board" into a peripheral slot of any computer.

Solving Chess

NP problems quickly exceed conventional computer capacity, but of exponential problems they are the easiest to solve. In an NP problem there are an exponential number of candidate solutions to be considered, but the correctness and cost of each candidate (a salesman tour, a circuit for a function, a mathematical proof) can be checked easily, a fact our time-loop NP computer exploits. But there exist harder exponential problems where evaluating a single candidate answer is itself an exponential problem.

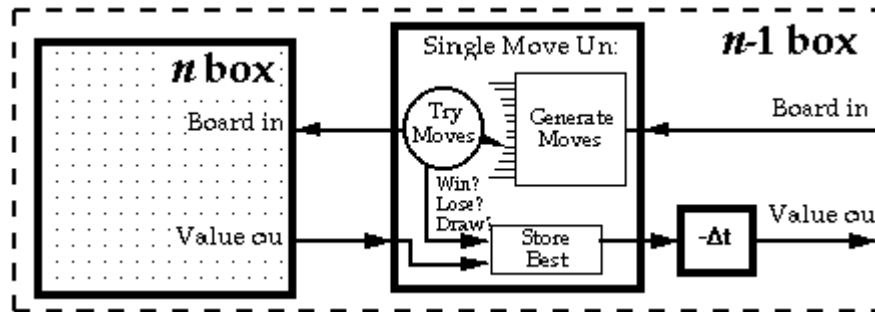
Finding the best move in a game like chess is an instance of such a problem. Your best move can be found by considering all possibilities for your move, then for each of those, all possible responses by the opponent, and all your possible responses to all of those, and so on, until a "tree" of all possible games has been mapped out. Then, working backwards from the "leaves" of this tree, i.e. the final moves of the games, all but the best moves can be pruned away. It is easy to evaluate possible last moves--from the point of view of the player whose turn it is, some are wins, some are losses and some result in draws, and no finer distinctions need be made. When all but the best last moves are discarded, the relative merits of second-to-last moves become apparent--each is given by the best last move that results from it. This pruning proceeds to earlier and earlier moves until the beginning of the game, and the best first move, is met. The chain of best moves from first to last is a perfect game of chess.

One can imagine a kind of generalized chess played on a variety of board sizes, with 8 by 8 giving the standard game. Finding a perfect game is easy on small boards, with few possible moves, but becomes enormously difficult as the board grows. The standard game is already so difficult that, despite known mathematical shortcuts, there is no hope that a conventional computer, even one using all the time and matter in the universe, could search the entire game tree. Today's chess computers examine a few levels of the tree, and use a formula to (very imperfectly) guess the value of the rest.

But time loop computers seem to greatly transcend the power of conventional machines. Can the approach used for NP problems be applied here? If, instead of city tours, the input to the computation box is the move sequence

for an entire game, the box could be arranged to stabilize only on sufficiently "good" games. But, unlike traveling salesman tours, games have only three values--win, lose or draw (for white, say). There seems no easy way to assign a number to indicate how nearly perfect a game is--and a perfect game itself is recognized only by considering and eliminating all other possible games.

There may be a devious way, however, to use negative delays to fold the massive tree search in time, in a fashion that makes the NP solver look positively pedestrian. This construction is going to resemble a mathematical induction. Suppose we have a circuit (call it the n box) that, given a particular chessboard position n moves into a game, is able to immediately tell us its value (win, lose or draw), as if it had searched the entire move tree from there. We could then build a box that provides the same services for a position $n - 1$ moves into the game by adding a "single move unit" that takes the $n - 1$ board, and, one by one, generates the possible next moves (typically about 30 of them), and feeds the resulting boards to the n box. It compares the n box results to one another, and chooses the best (for the player whose move it is). The time taken to do this is cancelled by a negative time delay element, and so the $n - 1$ box, like the n box, produces the value of a position as soon as the position is presented:



The n box is itself composed of a single move unit and an $n + 1$ box, and so on. The chain can be stopped eventually because each single move unit has a "short circuit" path that it uses when it encounters a position that, because of a king capture or excessive move count, is already a win, loss, or draw and so requires no further search. The whole machine is simply a chain of single move units and negative time delays as long as the longest game (a few hundred moves for standard chess), with the last unit set to permanently return a draw, signifying a stalemate.

Feed it the standard starting chess setup and the machine will immediately indicate whether the game is a forced win for white, black or a draw. To play (perfect) chess, give the machine the result of each of your possible moves, and choose one it finds best. But what is the machine up to in these deliberations? Having accepted time travel logic, the operation of the first single move unit is easy to intuit--it simply evaluates a few dozen possibilities and returns the answer, through the agency of a negative time delay, before its deliberations are complete. The second stage is more mysterious. Its input is changed by the first unit after it has delivered an answer, but while it is still doing the calculations for that answer! It is being asked to do several dozen different computations simultaneously. The third and subsequent units are required to be in even more bewildering states of superposition. The machine would certainly not work if its interior were being observed, as quantum observation collapses states of superposition to particular possibilities. Each single move unit would be seen to receive new boards to evaluate before it had completed the previous ones. Even weak observations (like hearing a whisper through a heavy wall), or indirect or delayed ones, would cause occasional collapses. To have any chance of giving correct answers, the machine must be totally protected from any peeking, perhaps by burying it in many layers of exotic shielding, or by placing it out of range of observation. Even then, its operation depends on aspects of reality hidden from present day experiments, such as the actual independent existence of the alternative worlds that appear in quantum mechanical calculations. Time will tell.

The non Computable

Exponential problems are big, but finite. Our universe may not be large or long lasting enough to solve serious instances of them without time travel, but it is possible to conceive of universes that are. But some problems are so hard that no bounded universe would be sufficient--their complexity is infinite. Kurt Gödel, who discovered rotating-universe time travel in general relativity, is much better known for shocking the mathematical community with his incompleteness theorems, showing that in any consistent and sufficiently interesting mathematical theory, there are unverifiable truths.

In the seventeenth century Pierre Fermat wrote that he had discovered a truly remarkable proof that there are no integer solutions to the equation $X^n + Y^n = Z^n$ when $X, Y, Z > 1$ and $n > 2$, but unfortunately the margin of the book he was annotating was too small to contain it. Subsequent generations of mathematicians have searched in vain for this (or any) proof to "Fermat's last theorem". Nor has a counterexample been found to prove it wrong. Most probably there was a subtle, unnoticed, flaw in Fermat's "proof". Today the theorem remains a good candidate for a Gödelian true but unprovable statement in arithmetic. If so, we won't be able to prove that it is unprovable, since such a proof would imply that no counterexample (providing a negative proof of the theorem) can be found, thus proving that no such counterexample exists, thus proving that the theorem is true, contradicting its unprovability.

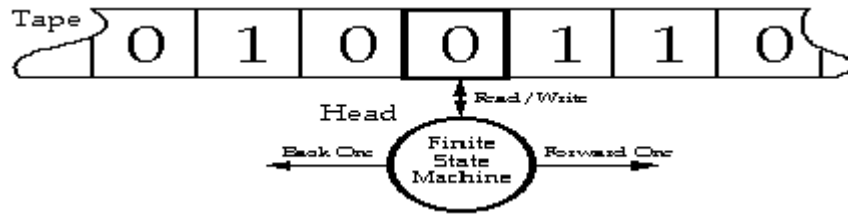
But Gödel's theorems hinge on the finiteness of proofs. We might try to evade them with time loops. Suppose, as in the NP machine example, we build an F box that takes as input a quartet of numbers $\{X, Y, Z, n\}$ and tests them to see if they constitute a counterexample to Fermat's conjecture. If so, the same numbers are presented at the output of the box. If not, the "next" quartet is generated and fed via a negative time delay back to the input. When the device is switched on, it immediately shows a counterexample to the conjecture, if such exists, proving the conjecture false. If, on the other hand, the box's signals hover in an intermediate state, the theorem must be true.

There's a problem, of course. A finitely sized machine can examine numbers only as wide as its signal paths. The time loop gives a lot of leverage, since, in an instant, it examines a number of cases exponential in the number of digits in these paths. But the number is still finite, and if the machine fails to find a counterexample, one can't be sure there isn't one that just happens to exceed its capacity. A search for any other kind of proof or disproof founders on this same Gödelian impasse--though a time loop machine can examine all alternatives of a given length, a wonderful proof may yet elude us because the machine we've built is too small to contain it.

In situations like Fermat's last theorem we may be willing to forgo an astronomically (or infinitely) large proof or counterexample if we are told by a trustworthy source that the statement is true or false--a compact answer that fits in a very small machine. The time computer techniques we've discussed thus far examine, in a short fixed period, a number of cases exponential in the size of the machine. It may be possible to substitute time for space, to build a small, continuously operating machine that examines an exponentially increasing number of cases as time passes. Take, for instance, a single move unit from the chess machine example, and connect its output to its input via a negative time delay. Observed, the system acts as a simple time loop, as in the NP example. But unobserved, the move sequencing might spawn multiple alternatives, just as in the chess solver, but of the original unit itself rather than units downstream. As time goes by, each copy repeatedly splits itself again, and the number of parallel machines grows exponentially. After a few hundred iterations, this single unit will have covered the same tree of alternative games as the extended chess machine. But by letting it run beyond that point, longer games can be examined. The answer is found by the peculiar procedure of inputting the initial chess board to the single move unit and peeking at its output value at the start of the computation, then tightly closing an observation-proof box around it. The answer you receive right away is trustworthy only if the machine is subsequently left to run, undisturbed and unobserved, for long enough.

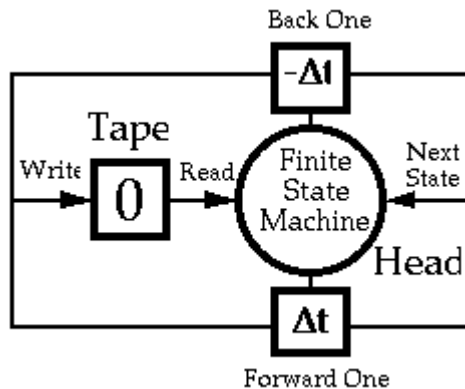
For more general problems, let's consider an abstract device conceived in the 1930s by British mathematician Alan Turing, inspired by David Hilbert's concept of "mechanizing" mathematics. A Turing Machine is a finite, usually simple, computer or "finite state machine" connected to a "read/write head" that moves back and forth on an indefinitely long tape. The machine proceeds in regular steps, at each reading a symbol on the square of the tape under the head, then, controlled by an unambiguous finite internal list of rules, writing a symbol in its place, and moving one square forward or back. By convention, the internal rulebook for any given Turing Machine is

fixed, but its tape can be initialized with an arbitrary string of symbols. Turing showed that such machines, generally using lots of tape to store and reference initial inputs, intermediate results and final answers, existed for every conceivable computation:



Turing also showed that there exist machines that interpret the initial contents of their tape as a rulebook of another Turing Machine, and proceed, slowly, to simulate this other machine. These are called "Universal" Turing Machines, for their ability to do any computation that another machine can do. A Universal Turing Machine is a good mathematical model for a digital computer, and Turing used the concept to prove "non-computability" theorems that are equivalent to (but more straightforward than) Gödel's "unprovability" results.

By appropriately initializing its tape, one can program a Universal Turing Machine, like a conventional computer, to do just about any computation, for instance to search through $\{X, Y, Z, n\}$ quartets looking for a counterexample to Fermat's Last Theorem, or through inference chains for a proof or disproof. A literal Turing Machine, though, is very slow--it spends most of its time crawling back and forth across increasingly long stretches of tape to get at widely separated bits of information. But it is amenable to a time-computer transformation. Imagine that the tape runs not in space, but in time. At any instant it is just a single square (probably a circuit) storing a symbol. Instead of moving left or right, the head sends a message to its past self through a negative delay, or to its future self through a positive delay:



Many questions could be posed about this design. A few have answers. Because the machine is created at a particular time, and does not exist before then, the "tape" ends abruptly in the backwards direction. This is not a problem because in any computation there are straightforward ways to "fold" tape usage to exclusively the forward portion. Interesting computations will use a lot of tape--i.e. will extend far into the future. But if the program copies its ultimate answer to the cell initially under the head, the answer will be immediately available. As with the the folded chess machine, one must start the "Temporal Turing Machine", quickly peek at its tape cell for the answer, and then seal it up, to let the computation run undisturbed, unobserved. But then how is the program installed on the tape, on cells corresponding to time periods when the machine is sealed? The easiest answer is to encode it in the finite state portion of the machine rather than the tape.

In its sealed box, the head effectively runs forward and backward in time, repeatedly overwriting cells, spawning multiple worlds ever faster. Yet the number of situations the machine can examine (different Fermat's Last Theorem quartets, for instance), can at most grow exponentially in the machine's running time. This is truly prodigious by conventional standards, but still quite finite. So here we are, as far out on a speculative limb as I'm willing to climb, and infinite problems remain hopelessly out of reach. It's time to quit.

Time Travel, Consciousness and Reality

The above constructions are no less strange when applied to "conventional" time-travel scenarios. Imagine that the computation boxes actually contain human beings--a person can quite plausibly evaluate the length of a traveling salesman tour, or enumerate the possible next moves in a chess game. Assuming, as before, that all observations must be logically consistent, what would a human, so embedded in a time loop, experience? When the NP machine produces a solution, its computation box evaluates only one case--a correct one, since quantum interference cancels others. To the person in the computation box, no less than to those outside, a correct answer appears, as if by magic, on the box's input--found correct, transcribed to the output, and relayed back in time, it also creates that input--a circular but logically consistent situation. On the other hand, there is little reason to believe that worlds containing incorrect solutions can be experienced, since such experiences would involve logical contradictions.

Things are not so straightforward for persons inside the chess machine--let's put a friend in each single move unit. The opening position unit feeds first moves one after another to the rest of machine (the "2" box), and examines the responses--no difficulty. But the second unit is required to be in a superposition of states, simultaneously examining each of the first move's alternatives. When made of dumb machinery, it was reasonable to simply seal it off from observation, to prevent spoiling the computation by collapsing the superposition to a single possibility. But an intelligent friend will probably need to be released sooner or later, and will retain a memory of the computation. In the strange logic of quantum mechanics, this memory, released from the box, constitutes an observation that retroactively collapses the computation. We could, of course, preserve the superposition and the computation by callously keeping the boxes with our friends permanently sealed, or dropping them down a black hole, or accelerating them away beyond observation. But then there will be no way to learn about their experiences. So, apparently, one can either exploit parallel worlds, or experience a single one of them, but not both. "All the world's a stage" wrote William Shakespeare, and apparently we players act in only one story at a time. But when we close our eyes and listen closely, we hear from the wings the echoes of other stories. Whether they really are other plays (with audiences?), or just sound effects by a clever stage manager, remains undecided.