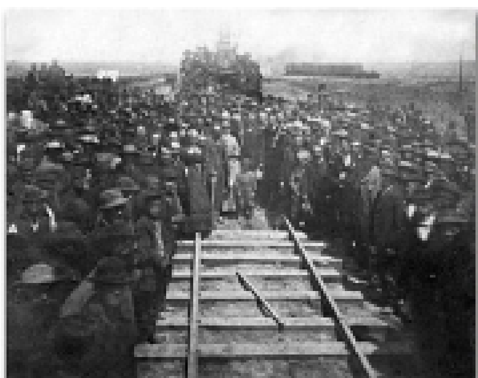
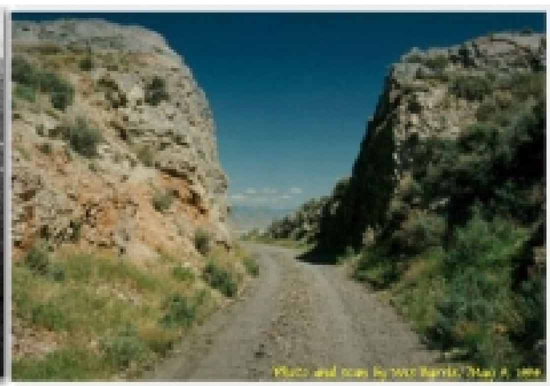
Information Technology: Will this, too, Pass?

Information technology seems to be the shaping force in society today. It's hard to imagine a world in which all those computers could disappear or fade to insignificance. Yet if history is any guide, all such technologies have a way of losing their dominance. In the last century it was the railroads that shaped the nation, but now it is hard to believe that Promontory Point, an unremarkable path northwest of Ogden, Utah, was long ago the setting for one of the most celebrated events in American history. It was here at 12:47pm on May 10, 1869 that the "golden spike" was driven home, connecting the Union Pacific Railway with the Central Pacific Railway to complete the last link of the transcontinental railway. If there had been television then, perhaps this event would have been likened to Neil Armstrong stepping onto the moon. But, of course, there wasn't, and instead the news was to be relayed to the waiting nation via the communications network of that day -- the telegraph.



The Last Rail is Laid
1869



Today

The ceremonial spike and hammer had been wired such that when they were struck together the "sound" would be immediately conveyed over the telegraph, which had completed its own continental joining some years earlier. However, it was said that both Governor Leland Stanford and Thomas Durant of the Union Pacific missed the spike with their swings. Nonetheless, the telegraph operator dutifully entered the message by hand, keying in three dots, signaling the single word "done."

It had been 25 years earlier that Samuel Morse had sent his famous message, "What hath God Wrought," from the Capitol in Washington, D.C., to Mount Clare Depot in Baltimore, thus commencing what was then considered the greatest information revolution ever. By the time the railways were joined, Western Union already had over 50,000 miles of cable and poles, and was by the standards of that time an economic giant. The electrification of the nation was just beginning. It would be another 9 years before the electrical light would be invented, and another four years before the first street in New York City would be illuminated. The telephone would be invented in six years. The three great infrastructures – electrical power, communications, and transportation – were then all in their infancy, but fertile with promise.

The railroad of that day had a transforming effect in industry that was similar to that we now ascribe to information technology. Commerce, which had been oriented north to south because of rivers and canals, was turned east to west. Chicago, which had a population of only 30,000 people in 1850, became the railway capitol of the nation and tripled its population both in that decade and the next. The radius of economic viability for wheat, the point at which the cost of transportation equaled the value of the wheat, expanded from 300 miles to 3000 miles. If there had been a phrase like "global village" in that day, people would have said that it was happening because of the railway.

Today the transforming power of the railways is long in the past. Promontory Point no longer seems like a crucial place. The golden spike itself was "undriven" as part of the war effort in 1943, and the point where the rail lines were joined is now a jogging trail. Transcontinental railways have had diminishing importance, and many have ceased operation. The telegraph has disappeared too, and AT&T has even removed the word “telegraph” from its legal name. Today it's all about information technology -- the force that will rewrite the rules of commerce, collapse the worlds of time and space, reinvent governments, and irrevocably change the sociology of the peoples of earth. That, of course, is what we say now. The question is whether this importance is transient, or whether information technology is so fundamentally different that its power will last into the millenium to come.

I asked a number of influential friends in the information technology business whether they thought that this current technology was fundamentally different than the transforming technologies of previous centuries. Everyone said that of course it was, and some implied that it was sacrilegious to even raise the issue. The conviction held by almost everyone in the field is that we're onto something that will change the world paradigm, not just for now, but for a long time into the future, and in ways that none of us is able to forecast.

So what is different? Most importantly, information technology seems almost infinitely expandable, not being bound by the physical laws of the universe that have constrained previous technologies. The problem with those railways was the intrinsic limitation of the technology. I imagine going back as a time traveler to Promontory

Point in 1869 to talk to people about the future of the railroad. People would tell me how wonderful and long-lasting the power of the railroads would be. Yes, I would agree, the railway would reshape the nation. It would be the dominant economic force for many years to come. But, I would say, look at that train. A century from now it would be pretty much the same. A hundred years of progress would perhaps triple the speed of the fastest trains. The strongest engines could pull more cars, and there would be maybe ten to a hundred times the track mileage, but that's it. What you see is basically what there is. There isn't anything else.

From my viewpoint as a time traveler in 1869 I would be quietly thinking that some decades later Ford would be building little compartments that propelled themselves without the need for tracks. Before long the Wright brothers would be testing contraptions that flew through the air. They wouldn’t be called winged trains, because the flying things would seem different enough to deserve their own name. The nation would become webbed with highways and clogged with traffic, and the skies would be filled with giant flying compartments. All these inventions would constitute a steady improvement in the speed, convenience, and cost-effectiveness of transportation. Still, transportation wouldn’t have improved all that much. Even today, looking at a lumbering Boeing 747 lifting its huge bulk into the welcoming sky, I wonder: how much better can transportation get? What comes after the airplane?

On the other hand, if they asked me about the future of the telegraph, I would shake my head in perplexity. A good idea, I would say, but far, far ahead of its proper time. The telegraph was digital, and being digital would be one of the three synergistic concepts that would constitute information technology. By itself, however, the telegraph could not be self-sustaining. What would be needed, and would not be available for another century, would be the other two concepts – the supporting technology of microelectronics and the overriding metaphysical idea of an information economy.

The telegraph, of course, was undone by the telephone. It was only a half dozen years after the golden spike was driven that Bell invented the telephone. His invention was basically the idea of analog -- that is, the transmitted voltage should be proportional to the air pressure from speech. In that era it was the right thing to do. After all, the world we experience appears to be analog. Things like time and distance, and quantities like voltages and sound pressures seem to be continuous in value, whereas the idea of bits – ones and zeros, or dots and dashes – is an artificial creation, seemingly unrepresentative of reality.

For over a century thereafter, the wires that marched across the nation would be connected to telephones, and the transmission of the voiced information would be analog. Not until the late 1950s would communication engineers begin to understand the intrinsic power of being digital. An analog wave experiences and accumulates imperfection from the distortions and noise that it encounters on its journey, and the wave is like Humpty-Dumpty – all the King's horses and all the King's men cannot put it back together again. There is no notion of perfection.

Bits, in contrast, have an inherent model of perfection – they can *only* be ones and zeros. Because of this constraint, they can be restored or regenerated whenever they deviate from these singular values. Bits are perfect and forever, whereas analog is messy and transient. The audio compact disc is a good example. You can play a disc for years and years, and every time the disc is played it will sound exactly the same. You can even draw a knife across the disc without hearing a single resultant hiccup in the sound. In most cases all of the bits corrupted by the cut will be restored by digital error-correcting codes. In contrast, many of us remember the vinyl long-playing records that sounded so wonderful when they were first played, but deteriorated continuously thereafter. And never mind that trick with the knife!

The underlying notion of the perfection of digital makes so much of information technology possible. Pictures can be manipulated, sound can be processed, and information can be manipulated endlessly without the worry of corruption. Any fuzziness or ambiguity in interpretation can be added on top of the stable digital foundation, to which we can always return. Written text is an example of fuzziness on top of digital stability. The code is the alphabet, which is a finite set of possible characters. Though we all might agree on which letters had been written, we might argue about the meaning of the text. Nature sets the perfect example in life itself – the instructions for replication are transmitted in the four-letter digital code of DNA. At the very base level we find the indestructibility of digits, without which life could not be propagated and sustained.

Is digital here to stay? Or could a budding young Alexander Graham Bell make some invention that convinces us to go back to analog? While there are compelling reasons why being digital is a foundation of the information age, it also true that digital implementation perfectly compliments the microelectronics technology that is the building block of choice today. This is a stretch, but I suppose I could imagine microelectronics in the future being replaced by quantum devices or biological computers that were intrinsically analog in nature. People might say, "Remember when computers only gave you a single answer, like everything was so crystal clear, and they didn't realize that life is fuzzy and full of uncertainty?" I suppose, but I don't think so. I think digital is here to stay.

The digital revolution started before microelectronics had been developed, but it is hard to imagine that it would have gotten very far without the integrated circuit. The first digital computer, ENIAC, was built from vacuum tubes in 1946 at the University of Pennsylvania, nearly at the same time that the transistor was being invented at Bell Labs. But digital computers were unwieldy monsters, and for the decade of the 1950s it was analog computers that achieved popularity among scientists and engineers. Then in 1958 a young engineer named Jack Kilby joined Texas Instruments in Dallas. Being new to the company, he was not entitled to the vacation that everyone else took in the company shutdown that year. Working alone in the laboratory, he fabricated the first integrated circuit, combining several transistors on a single substrate, and thus beginning a remarkable march of progress in microelectronics that today reaches towards astronomical proportions.

The integrated circuit is now such an everyday miracle that we take it for granted. I can remember when portable radios used to proclaim in big letters, "7 Transistors!" as if this were much better than the mere 6-transistor imitators. Today you would never see an ad saying that a computer features a microprocessor with, say, ten million transistors. Like, nobody cares. But I watch that number grow like the number of hamburgers that McDonald's sells, because it is the fuel of the digital revolution. If it were to stop growing, ominous things would happen.

In 1956 Gordon Moore, a founder of Intel Corporation, made the observation that the feature sizes (the width of the wires and sizes of the device structures) in microelectronics circuitry were growing smaller with time at an exponential rate. In other words, we were learning how to make transistors smaller and smaller at a rate akin to compound interest. Integrated circuits were doubling their cost-effectiveness every 18 months. Every year and a half, for the same cost, we could produce chips that had twice as many transistors. This came to be known as Moore's Law, and for twenty-five years it has remained almost exactly correct.

Moore's Law is the furnace in the basement of the information revolution. It burns hotter and hotter, churning out digital processing power at an ever-increasing rate. Computers get ever more powerful, complexity continues to increase, and time scales get ever shorter -- all consequences of this irresistible march of technology. Railroad technology had nothing similar. Perhaps no equivalent technology force has ever existed previously.

That the power of digital electronics is increasing exponentially is difficult to comprehend for most people. We all tend to think in linear terms -- everything becomes a straight line. The idea of doubling upon doubling upon doubling is fundamental to what is happening in electronics, but it is not at all intuitive, either to engineers and scientists or to decision-makers in boardrooms. This kind of exponential growth can be envisioned by recalling the old story of the king, the peasant, and the chessboard. In this fable the peasant has done a favor for the king and is asked what he would like for a reward. The peasant says that he would wish simply to be given a single grain of rice on the first square of his chessboard, and then twice as many grains on each succeeding square. Since this sounds simple, the king agrees.

How much rice does this require?   I discovered that one university has a physics experiment based on this fable to give students an intuitive understanding of exponentiation. The students are given rice and a chessboard to see for themselves how quickly exponentials can increase. They discover that at the beginning of the board very little rice is required. The first 18 to 20 squares of the board can be handled easily using the rice in a small wastebasket. The next couple of squares need a large wastebasket. Squares 23 through 27 take an area of rice about the size of a large lecture table. Squares 28 through 37 take up about a room. To get to the last square -- the 64th -- requires a number of grains represented by a 2 followed by nineteen zeros - variously estimated at requiring the entire area of earth to produce, weighing 100 billion tons, filling one million large ships, or one billion swimming pools.

This is the way exponentials work. At first they are easy, but later they become overwhelming. Moore's law says there will be exponential progress in microelectronics and that doublings will occur every year and a half. Since the invention of the transistor there have been about 32 of the eighteen-month doublings of the technology as predicted by Moore - the first half of the chessboard. Should this law continue, what overwhelming implications await us now as we begin the second half of the board? Or is it impossible that this growth can continue?

Moore's Law is an incredible phenomenon. Why does it work so well? After all, this is not a law of physics, or a law that has been derived mathematically from basic principles. It is merely an observation of progress by an industry – an observation, however, that has been an accurate predictor for a quarter of a century. In spite of the fact that this law is the most important technological and economic imperative of our time, there is no accepted explanation for its validity.

Gordon Moore himself has suggested that his law is simply a self-fulfilling prophecy. Since every competing company in the industry "knows" what progress is required to keep pace, it pours all necessary resources into the pursuit of the growth of Moore's Law. This has required an ever-increasing investment, as the cost of fabrication plants necessary to produce smaller and smaller circuitry has escalated continuously to over a billion dollars per facility. Or perhaps this is simply an instance of positive feedback. As the electronics industry grows to a larger proportion of the economy, more people become engaged in that field and consequently there is a faster rate of progress, which in turn grows the sector in the industry. Whatever the reason behind Moore's Law, it is more likely to lie in the domains of economics or sociology, than in provable mathematics or physics, as is the case of the more traditional "laws" of nature.

Aside from the puzzlement of why Moore's Law exists at all, there is the puzzling question of why the period of doubling is 18 months. {{this doesn't seem important, it seems like a coincidence. There can't be an inherent law of nature that governs product cycles!This rate of progress seems to be critically balanced on the knife-edge of just the amount of technological disruption that our society can tolerate. I sometimes think that Moore's eighteen months is akin to the Hubble constant for the expansion of the universe. Just as astronomers worry about the ultimate fate of the universe – whether it will collapse or expand indefinitely – so should engineers and economists worry about the future fate of Moore's expansion constant. Suppose, for example, that Moore's Law ran much faster, so that doublings occurred every 6 months. Or suppose that it was much slower. What would happen if Moore's Law suddenly stopped?

We've grown so used to the steady progression of cheaper, faster, and better electronics that it has become a way of life. We know that when we buy a computer, it will be obsolete in about two years. We know that our old computer will be worthless, and that even charities won't accept it. We know that new software may not run on our old machine, and that its memory capacity will be insufficient for most new purposes. We know that there will be new accessories and add-on cards that can't be added to our old machine. We know all this, but we accept it as a consequence of the accelerating pace of information technology.

One very good reason why the eighteen-month period of doubling is not disruptive is a law of economics that helps counterbalance the chaos implied by Moore's upheavals – it is the principle of increasing returns, or the idea of demand-side economies of scale. In other words, the more people who share an application, the greater is the value to each user. If you have your own individual word-processor that produces files that are incompatible with anyone else's system, it has very little value. If you have a processor that is owned by few other people, than you will find it difficult to buy software for it. For this reason the information technology market often has the characteristic of "locking in" to certain popular products, such as Microsoft Word or the PC platform.

The good news is that the economic imperative of increasing returns forces standards and common platforms that survive the turmoil of Moore's Law. The bad news is that it is often hard to innovate new products and services that lie outside the current platform. For example, the Internet today has been characterized as old protocols, surviving essentially from their original design about 30 years ago, running on new machines that have been made continuously more powerful by Moore's Law. We can't change the protocols, because the cost and disruption to the hundred million users would be enormous.

Now imagine a world in which Moore's Law speeds up. The computer you buy would be obsolete before you got it hooked up. There would be no PC or Apple standard platform, because the industry would be in constant turmoil. Every piece of commercial software would be specific to an individual processor, and there would be flea markets where people would search for software suitable for the computer they had bought only a few months before. Chaos would abound.

But suppose instead that Moore's Law simply runs out of gas. Suppose we reach the limits of microelectronics where quantum effects limit the ability to make circuits any smaller. What would the world of information technology be like? Perhaps computers would be like toasters, keeping the same functionality over decades. Competition between computer manufacturers would be mainly in style, color, and brand. Maybe there would be mahogany computers to fit in with living room décor. Used computers would hold their value, and programs would stay constant over years. Intel, Microsoft, and other bastions of the tech market sector would have to find new business models.

Could it happen? Could computers stop getting better? It may sound as if an end to computer improvement is inevitable, yet there is scientific disagreement on this possibility. Over the last two decades many scientific papers have been written predicting the end of Moore's Law because of the constraints imposed by some physical law. So far they have all been wrong. Every time a physical limit has been approached by the technology, a way around that limit has been discovered. However, similar papers are still being written, and they are looking more and more credible as the size of circuits descends into the world of quantum mechanics and atomic distances. While there is somewhat of a consensus that Moore's Law could continue for another decade, further doublings become more and more problematical. Remember the fable about the rice and the chessboard and the awesome consequences of unlimited exponential growth. Surely there is a limit, and it may not be that far away. The likelihood is not so much that Moore's Law would stop completely, but that it would begin to slow down.

Even as we approach possible physical limits to the technology, however, there is a huge industry that will resist the slowing of technological growth with all of its collective strength. The beat must go on, somehow, someway. Material systems other than silicon that promise smaller circuit sizes are being explored. Some researchers are betting on quantum or biological computing. Meanwhile, research on the algorithms that process information continues unabated on its own Moore's Law curve. In a number of important areas like image compression, more progress has been made through mathematical ingenuity than brute force application of faster computing.

Going forward, however, the biggest potential for continuing improvement lies in networking and distributed computing. Today there are two parallel developments taking shape – the evolution towards higher speed networks, and the invasion of a myriad of embedded, invisible, and connected computers. In networking we know that it is possible to increase the capacity of optical fibers a thousand fold, which will result in a veritable sea of capacity at the heart of the Internet. Technical obstacles to highspeed access to this sea are starting to be overcome, driven by a plethora of competing alternatives, including cable, telephone, wireless, and satellite. The demand for high speed is currently being fueled by audio, but video will soon follow. Within a few years megabit access will be common, followed shortly by tens of megabits at each home. In the sea beyond a kind of distributed intelligence may take shape.

At the periphery of the network tiny inexpensive devices with wireless Internet connectivity are likely to cover the earth. We are imagining that every appliance, every light switch, every thermostat, every automobile, etc., will be networked. Everything will be connected to everything else. Networked cameras and sensors will be everywhere. The military will develop intelligent cockroach-like devices that crawl under doors and look and listen. The likelihood exists that everything will be seen, and that everything will be known. We may well approach what author David Brin has called the transparent society.

Because of these trends, it is probable that a couple of decades from now the computer as we know it today – an isolated, complex, steel box requiring hour upon hour of frustrating maintenance by its user – will be a collector's item. Instead we will all be embedded in a grid of seamless computing. We may even be unaware of its existence. People will look at computers in museums or in old movies and say how quaint it must have been. Unlike the railway, which after centuries still has its same form and function, computers will have undergone a metamorphosis into a contiguous invisible utility. They will have moved into a new existence entirely. Hopefully, they will take us with them.

After being digital and the infrastructure of microelectronics, the third aspect of the current evolution is the idea of an information economy. Personally, I never cease to be amazed at how this new economy functions. When I was a youth the only adults I saw working were the carpenters building the new houses on my suburban street and the farm workers in the nearby fields. I grew to believe that was what adults did – they made things and they grew things. One day as an adult myself I came to the sudden realization that I did know a single person who made or grew anything. Everyone that I knew had rather mysterious jobs where they moved information or money around. Some of them had grown very wealthy doing these nebulous things. Sometimes I am wryly amused by this information economy. I myself am so far removed from any physical contribution to society that I have a sense of unreality. How do we all get away with it? We have all moved into the land of bits, leaving mostly machines to be the caretakers in the land of atoms.

The image of the golden spike being struck, enabling heavy, fire-breathing engines to clatter over iron tracks seems to epitomize the old world of atoms. Recently I waited on the platform of a railroad in a small Japanese city. I saw in the distance an express bullet train approaching the station. The speck grew quietly larger, until suddenly with something akin to a sonic boom it burst through the small station. The earth shook beneath my feet, and the vacuum and shock of its passing left me breathless. What awesome, visceral power!   But this is the world of atoms, and it is bound by the limits of the physical world -- by energy and mass, by the strength of materials and the depletion of natural resources. The world of bits, in contrast, seems to glide by in an ethereal serenity.

Information in itself weighs nothing, takes no space, and abides by no physical law other than possibly the speed of light. Moreover, information is non-rivalrous. Unlike a physical good, information can be given away, yet still retained. It can also be copied perfectly at near zero cost. It can be created from nothing, uses up no natural resources, and once created is virtually indestructible. We are still trying to cope with these properties as they clash with our traditional laws and business models in industries such as publishing and music.

In Norman Augustine's book, *Augustine's Laws*, he observes that nearly all airplanes cost the same on a per pound basis.   Tongue-in-cheek, he says that the industry must have been stalled for new revenues after the development of the heaviest planes, the Boeing 747 and the military transport C-5A. In order to increase the expense of their product they needed to find a new substance that could be added to airplanes that was enormously expensive, yet weightless. Fortunately, such a substance was found -- it was called software. Thereafter the cost of planes could increase indefinitely without regard to their lift capacity.

Having escaped the bounds of physical laws, information technology can continue to expand in unforeseen directions. The World Wide Web came as a surprise to nearly everyone in the business, even though it appears obvious in retrospect. It was a great social invention, based on old and available technology. If fact, nearly all the progress made today in the Internet world is based upon new business models and sociology, rather than technology. I often hear the expression "Technology isn't the problem." According to some people, we have enough already to last for a long time to come.

What comes after the information age? I have no idea. But as I walk the halls of industry, passing by cubicle after cubicle containing roboticized humans staring at CRT screens, I think this can't last. This strange amalgam of human, screen, keyboard and mouse must surely be a passing phenomenon. Perhaps in three or four decades one of my grandchildren will be able to say, "I don't know a single person who deals with information." Maybe then people will only deal with meta-information, or policy. Maybe we will rise above the land of bits into the land of ideas and dreams. Whatever there is, it will surely be different. But because information technology has escaped the bounds of physical limitations, because it is supported by microelectronics powered by the compound interest of Moore's Law, and because of the lasting beauty of being digital, I'm betting that it will endure in some form through the next century.

Robert W. Lucky