Virtual Molecular Reality
Prof. Marvin Minsky MIT Media Lab
Advisor, Foresight Institute

(In *Prospects in Nanotechnology,* Krummenacker and Lewis,, Eds., Wiley 1995)

Introduction by Dr. K. Eric Drexler

Our speaker tells me that for today at least he prefers not to be introduced as a co-founder of the field of artificial intelligence and of the MIT Artificial Intelligence Laboratory. For today he would prefer to be known as the inventor of the confocal scanning optical microscope. So, to speak on computation in the era of molecular manufacturing—which is a broad subject, and I am sure that we will find that the boundaries of this field are hard to place limits on—I would like to introduce Marvin Minsky.

1 Tinkertoy, Meccano, and Erector Sets

The great thing about being a professor at MIT is that you get to have the best teachers, like Eric Drexler. They are called "students," and if you pay attention to them very carefully, you can learn a great deal and even get credit for some of their work. Yes, I did sign his thesis on nanosystems—but if I would have tried to write one, he would not have signed mine.

According to my family, I was proficient with construction toys at an early age. Once, I built a structure of Tinkertoy rods that reached to the top of a hotel lobby. The adults tried to figure out how a person so small could build something so big. My problem was to figure out how people so large could think so small.

If you remember Tinkertoys, with its wooden rods and spools, you may be showing your age. They still exist, but, nowadays, LEGO is more popular. This I regret, because I feel that LEGO blocks might be a bit too Cartesian (that is, right-angle oriented) to encourage building more versatile structures. In recent years, LEGO has added hinges and other new parts so that you can join things with more diverse angles, but this makes the system less elegant.

The important thing about Tinkertoy—and other such construction sets—is not merely that they help you to understand structures, space, and geometry. They also convey a profound idea that could help children to understand certain other important domains. Such toys teach you the concept of a "Universal Set of Operations," the idea that a limited set of primitive operations can generate an infinite and rich variety of composite structures. For example, this is why chemistry is so rich, although it, too, like Tinkertoy, starts out with nothing more than a few kinds of atoms and a few kinds of bonds. As another example, this is why mathematics is so rich, although it starts out with nothing more than a few axioms and a few rules of inference. And this is same reason why Drexler's approach to eutactic nanotechnology is so rich.

There are some, I am sure, who would object that these elegant toys might create for the child a sterilized reality. "Better," those critics might maintain, "to give the child more realistic, less ideal materials. Give them things like clay and paint, things you can squeeze and smear, so their imaginations will be less constrained." There is something to be said for this—and I, myself spent a considerable amount of time trying to build with mud and sand. I suspect, however, that with those "more natural" materials, one can never go very far beyond certain levels of complexity.

I certainly found that to be the case when trying to make things by cutting and joining "continuous" materials. Whenever I would make things with wood, saw, and nails, the parts never fit quite properly, and the errors would accumulate. Thus, when I tried to build with wood, each piece would end up out-of-size by at least a sixteenth of an inch, because of my forgetting the width of the saw. It was only in much later years, when shaping steel with milling machines, that this kind of error quite disappeared. Instead, the piece I was working on would end up at precisely the desired dimension to within one ten-thousandth of an inch—but sometimes with an error of precisely plus or minus one inch! This is because, as machinists know, the most common mistake is to miscount the turns of the coarse-feed crank.

My point is that those construction-sets provided you with a "virtual reality"—an artificial universe in which things worked exactly as they were supposed to work. When you build things with Tinkertoy, Erector, or Meccano sets, a proper conception yields things that perform—and if they don't, then your scheme was wrong. You should not need to file or shim or bend those parts to make them fit—the way that everything else is done. Two objects assembled with those kinds of parts are exactly the same, or clearly not—precisely as it is with molecules, mathematics, or programming.

You know how, in the "real" world, things scarcely ever work very well? It does not help when adults explain, "It is not important to reach your goal. It is only important to do your best. It is the journey that matters, and not where you get." That is not wisdom, in my view, but only the sour-grape excuse of broken spirited mortals. Perhaps this is why I am so in love with nanotechnology. It is only partly because of nanotechnology's elegant new ideas and possibilities. Perhaps my love is more because it brings me back to building things in a perfect world in which everything works as it ought to work—a world in which ideals become achievable.

2 Quantum certainty

In espousing nanotechnology, both Eric and I have had problems with the academic establishment. Many otherwise excellent colleagues have objected that nanotechnology might be impractical because of quantum uncertainty. "You cannot be sure where the atoms are," they would complain, grumbling about the uncertainty principle, tunneling, and strange correlations. In Nanosystems [1], Drexler treated this carefully by calculating many such uncertainties, and showed that those skeptics were usually wrong. At ordinary temperatures, the thermal effects tend to dominate the quantum effects—and tend not to be very serious. In the rod-logic structures that he designed (without any redundancy), the most serious problem is damage from cosmic rays-and that would happen classically, too, if exposed to such energetic particles. Chemical structures like DNA

can be stable for many millions of years, if there are no other destructive processes.

In any case, I sometimes suspect that these objections are unconsciously based on a peculiar bit of philosophy that has insinuated itself into our popular intellectual culture. We see it in statements such as the following:

"In the world of classical mechanics, everything worked like clockwork, with deterministic certainty. Then quantum theory changed our view, so that now we understand that all is uncertain and indeterminate."

Almost everyone accepts this view of history. However, now I will explain why one should take an opposite view. Uncertainty was inherent in the classical view, while only quantum theory showed that things could be depended upon.

Let us explain this seeming paradox. Before quantum mechanics, the dominant idea was that matter was made of particles that interacted through inverse example, as described by Newtonian physics, that has a heavy object in the middle, and several lighter objects surrounding it. Such small systems appear to be rather badly behaved—and in the particular case of the solar system in which we live, no one has been able to show that it is stable. For example, Gerry Sussman (another "student" of mine) and Jack Wisdom at MIT seem to have shown that the orbit of Pluto is chaotic, and, so far as it is known now, Pluto may eventually get thrown out of the solar system. We Earth-people might not consider this a serious loss. However, consider that our largest planet, Jupiter, has enough angular momentum that, given suitable coupling, it could hurl Earth itself into outer space. I do not know if the Jovians would consider this a significant loss.

Thus, solar systems are unstable things. The same very likely would be true for atoms and molecules if they were similar, if their nuclei and electrons interacted by those classical laws. Even if each such atom by itself happened to be stable, when any two such atoms approached one another, the electron orbits inside them would soon be perturbed, and one or both atoms would soon break up. In the world of quantum theory, however, each atom is stable—completely unchanged—until there occurs a transition jump. The result is that we can have molecules with covalent bonds, in which the electrons maintain precisely at specific energy levels for billions of years.

Thus, contrary to what our science teachers tell our kids, it was in that old Newtonian World that almost everything would be unstable and indeterminate, whereas it is quantum mechanics that makes possible chemistry, life, and nanotechnology. It is because of quantum states that you can remember what you had for breakfast. This is because the new neural connections made in your brain can persist throughout your day. When something more important occurs, you will remember it as long as you live. Everything that we can depend upon exists because—it needs a name—because of "Quantum Certainty."

So, the next time you hear philosophers or physicists explaining that our world is based on statistical uncertainty, ask them if they realize that only Quantum Certainty makes anything we know persist. Is there any way to do the same within the so-called deterministic world of classical physics? I do not think I know the answer to that. It might be a good research topic.

3 Missteps toward micromachinery

When I was a child (a state that persisted for many years), I was deeply intrigued with the prospects of making things small. Then, in my college days, I talked with Tom Etter and Rollo Silver about the idea of making a tiny machine shop, with miniature lathes and milling machines. The idea was to use these tools to build minimanipulators that could work on a ten-times-smaller scale. (We all were partly inspired by Robert Heinlein's prophetic 1950 novel Waldo, whose inventor hero had a disease that made his muscles extremely weak. Fortunately, he was wealthy enough to have a satellite built so that he could live in zero gravity. Because he felt so isolated there, he invented teleoperators so that he could do what he wanted back down on Earth.)

Anyway, once you had made those smaller manipulators, you would use them to build another machine shop yet 10 times smaller, and so forth. Each such step might need some new technology (for example, because the previous generation of bearings and motors might not work well at the next smaller scale). We assumed that each such set of problems might be overcome in a decade of work. Now observe that a nanometer is only $10^{**}9$ meters—that is, a mere 10 of those decade-steps away. If you started today, you could easily assemble, with off-the-shelf parts, a cubic-meter machine shop. So, if all went well, in just 80 years you would have an atomic-scale assembler. We have not quite made that much progress in the 40-odd years that have passed since then, but we might be able to make up that time by discarding all those intermediate steps and going straight to eutactic technology. No one yet has demonstrated a foolproof way to build the first such assembler, but that is surely only a matter of time.

Early in the 1950s, Etter and Silver also conceived of "computing cloth"—an idea inspired by neon bulbs. Their scheme was to use closely spaced electrodes to move tiny plasma discharges inside a rare-gas atmosphere. (It requires a high voltage to start an arc, but lower voltages can make the arc move from one location to another, and they showed how to make ANDs, ORs, and Inverters this way.) The goal was to weave all the circuits into a microscopic insulated-wire cloth, while creating all the logic gates simply by removing the insulation at appropriate points. Etter and the Silver equipped the lab with vacuum equipment and glassblowing stuff, and eventually built a non-so-small shift register that used these gates. This was long before the earliest integrated circuits. Unfortunately they could not raise enough capital, so the project was abandoned.

Not long after that, Silver and I continued to dream about how to economically "stamp out" computers. We actually built some small hydraulic computer elements, by making millimeter-wide grooves and holes in multiple layers of plastic sheets, and placing little rods or balls in some of those grooves. When the assembly was pressed together and connected to a water supply, it became a computer, with circuits that worked at 30 Hz, powered by a 3-inch high column of water. Again, no one seemed interested in our little "hydroflip computers," although other forms of hydraulic logic were becoming popular for designing radiation-proof machinery.

Also, in 1955, with help from Ned Feder, I developed an elegant electrically controlled micromanipulator with

which you could easily write your name in submicron sized letters. It even had a way to estimate the force at its tip by vibrating the probe and viewing the amount of deflection. It simply never occurred to me to try making it sensitive enough to resolve individual atoms—and, therefore, I foolishly failed to invent the Atomic Force Microscope.

4 What to do next?

I like Drexler's schemes for making machines that use nothing but latches and rods—like making computers of Tinkertoy. Indeed, Danny Hillis and a group of other students made one of those once. This is stuff that clearly could work, and that even a child could understand. If at any stage you encounter a bug, you could always scale the whole thing up. I have heard people point out that we could make faster computers by using electrical circuits instead of rod logic, but this would surely lead to serious new problems. Strongly bound atoms do not jump around much (though we better be careful of hydrogen), whereas electrons might tunnel all over the place. If we decide to take that route, we must select some functional units to mass-produce.

The alternative to assembling parts is to develop a more integrated technology, through which a system is designed, simulated, and then fabricated in a nonmodular fashion. In the early days of "small-scale integration," computers were made by mounting separate logic modules onto a large, planar "mother board." The advantage of this was to make it easy to replace failed components. An amusing aspect of this was that by the mid-1970s, the price of small-scale logic chips had dropped down to 15 cents or so, but the sockets to hold them cost 5 times more—and often the connectors between the boards ended up costing more than the components themselves. When we begin to build our first nanomachines, will we begin with small modules to be assembled later, or should we try to do everything at once, by making the whole machine in one piece?

5 The skeptics

It was fascinating to see how skeptical people were, over the past decade, about whether nanomachines could be possible. I think that this era is close to its end, although there will always be those who make their livings (or at least their reputations) by being skeptical. In the field of artificial intelligence (AI), there are always one or two such critics, but usually not many more. Perhaps this is because there is room for only a few best-selling books in this area. In any case, it seems quite strange for anyone to argue that you cannot build powerful (but microscopic) machinery, considering that our very own cells prove that such machines can indeed exist. And then, if you look inside those cells, you will find smaller machines that cause disease. Most arguments against nanotechnology are arguments against life itself.

We see the same view (about AI) when skeptics assert that no matter what else we make our machines do, there is no way we could ever make them be conscious, or sentient, or anything like that. However, when you ask those critics precisely why this should be a problem—that is, what aspect of consciousness is beyond a machine—the answers that come back are almost identical with the arguments, a century ago, against molecular biology. "No mere arrangement of chemicals could possibly become alive." "Why not?" "Because you cannot reduce the vital spirit to anything so understandable." When this exchange is translated into the domain of AI, the phrase "vital spirit" is replaced by the term "consciousness." When you ask for more details, the discussion becomes exceedingly vague, and the skeptics cannot seem to tell you what they mean—except that machines cannot be "self-aware."

So, then, try asking these skeptics to what extent, and in what senses, are they aware of their own selves. First they will answer, "Clearly, I know that I am here." They could say that just as well, about the presence of their shoes. So, now ask, "Please tell me what is happening inside your own mind? How did you choose the words you just said? How did you put them in sentences? How did your brain recognize what I said? How did it know what I meant by them? And tell me, how did your brain decide which shoe to put on first today?" No human can answer any of these, because we simply have no access to how we choose one thought or expression over another, or why we first put on the left shoe. Whatever consciousness might be, it simply does not do the things that all those skeptics claim it does.

So, it seems to me that most of those critics go wrong at a truly critical point. If you think of consciousness as being aware of the processes that underlie your ability to think, then there is scarcely any reason to suppose that people are conscious at all. (If they were, we would not need psychologists.) Now, contrast this with an AI program running under a LISP interpreter that has its functional "trace" feature turned on. Then, whatever that machine might happen to do, you can always interrupt it by asking, "Why did you do that?" The machine is capable, in principle, of returning descriptions of how each decision was made, at least until it runs completely out of short-term (program-stack) memory.

My guess is that the most significant difference between the ordinary chimpanzees and the dressed-up chimpanzees is that, at various times in the past few million years ago, our brains augmented their capacity for "short-term pushdown" memory. That made it possible for certain parts of the brain to "replay" descriptions of what happened in other parts of the brain, and these enable us better to keep track of things, to analyze which thought processes had been most effective, and to remember methods that went wrong and still have some chance to straighten them out. What I am saying is that naturally, consciousness will remain a mystery until we have a good theory of it—and those critics do not have any such theories. It was the same, last century, when "self-reproducing" seemed obviously beyond what any machine could do. Today, we know dozens of ways, and all of them seem trivial.

6 Gauging progress

One kind of complaint about AI does have a little more merit. Many people assert that in the 1960s, AI seemed to be advancing rapidly, but it no longer is keeping its promises. The work appears to have slowed down so much that there has to be something fatally wrong. There are several answers to this.

The simplest answer is that we solved many easy problems in the first few years. Now we face harder problems, so naturally progress will tend to be slower. Certainly this is partly true, but not nearly as much as

those critics perceive. This is because, in every field, progress in the past 10 years rarely seems so fast as before, simply because you cannot yet tell which recent work was important. Certainly, during the 1980s, many of those who considered nanotechnology to be a science-fiction pipe dream were wholly and simply unaware of progress already underway in such domains as building artificial catalysts, and improving atomic-force scanning microscopy.

7 Discussion

AUDIENCE: Do you think that we are extending human capacity through things like artificial intelligence? Are we going to be a new form of life?

MINSKY: Yes, indeed. I find it appalling how many people are willing to tolerate the bad deal that they have been given. We ought to be more insistent about improving both our brains and our bodies. In regard to our brains, we have serious bugs. Isn't it amusing that the person you are talking to can remember a local phone number because it has only seven digits? If you add just three more digits of area code, your listener goes running for pencil and paper. You would think that an animal with trillions of synapses would be able to do much better than that, and this is one thing that we ought to improve.

I find it even more annoying that we have to live only a hundred years just because of a few evolutionary mistakes, such as the way that our hearts are fed. Why must we submit to those quadruple bypass operations? Simply because a half billion years ago, the prevertebrate heart was small enough to take fuel and oxygen by diffusion from the blood passing through it. When that got more difficult, our ancestors chose a simple, quick fix: just extend a little branch from that nearby artery into the myocardium. It might have required a few more genes to do this in parallel at many sites inside the pericardium. That is not what happened, because evolution had no plan for what would happen later, and this mistake led to premature death for many of us. Evolution has no direct way to remedy architectural errors made eons ago. When we design new forms for ourselves, we will describe our intentions along with the plans.

Now, the heart is little more than a pump. The brain, however, is more of a challenge. Certainly, nanosurgery will correct most diseases and the causes of aging, too. Eventually, we will want to remedy the inborn deficiencies of the brain. In this domain, we are surely approaching what the great science fiction writer, Vernor Vinge, calls a "period of singularity." Within a historically very brief time, we will have to face decisions about the homes we shall make for transplanting our minds. There is little we can do today, except to preserve ourselves cryogenically, in wait for the time when it will be routine to run nanoprobes into the brain, download all the knowledge therein, and later load it back again into a more capacious and reliable "brain."

I will go no further in this direction for fear of scaring away the potential nanotechnology research sponsors in this audience. For those who are concerned that this domain seems too good to be true, I suppose that the same was often said about the prospects of electricity. I am certain that this field will dominate the next century.

AUDIENCE: Do you anticipate the development of a hacker culture for nanotechnology?

MINSKY: There are hackers, and there are crackers. There are glorious prospects and there are dreadful prospects. It seems to me that a way must be found to keep things open enough so that we can catch malicious people before they can do anything too bad. Accomplishing that will not be easy. We might have to give up our privacy. There are terrible things in the universe. Quasars, for example, appear to be galaxies that exploded because something bad happened there. I wonder how many of those were science-fair projects that got out of hand.

Slightly revised 05/05/2008

[1] Drexler, K.Eric, *Nanosystems: Molecular Machinery, Manufacturing, and Computation.* New York: John Wiley & Sons, 1992.