Afterword to Vernor Vinge's novel, *"True Names".*

Marvin Minsky, October 1, 1984

(This web-version was considerably revised.)

In real life, you often have to deal with things you don't completely understand. You drive a car, not knowing how its engine works. You ride as passenger in someone else's car, not knowing how that driver works. And strangest of all, you sometimes drive yourself to work, not knowing how you work, yourself.

Then, how do we manage to cope with things we don't understand? And, how do we ever understand anything in the first place? Almost always, I think, by using analogies—by pretending that each alien thing we see resembles something we already know. Whenever an object's internal workings are too strange, complicated, or unknown to deal with directly, we try to extract what parts of its behavior seem familiar—and then represent them by familiar symbols—that is, be the names of things we already know which we think behave in similar ways. That way, we make each novelty at least appear to be like something we already know from our own pasts. It is a great idea, that use of symbols. It lets our minds transform the strange into the commonplace. It is the same with names.

For example, suppose that some architect invented a new way to go from one place to another: a device which serves in some respects the normal functions of a door, but one whose form and mechanism is so entirely outside our past experience that, to see it, we'd never of think of it as a door, nor guess what purpose to use it for. No matter: just superimpose, on its exterior, some decoration that reminds one of a door. We could clothe it in a rectangular shape, or add to it a waist-high knob, or a push-plate, or a sign lettered "EXIT" in red and white, or do whatever else may seem appropriate–and every visitor will know, without a conscious thought, that pseudo-portal's purpose–and how to make it do its job.

At first this may seem mere trickery; after all, this new invention, which we decorate to look like a door, is not really a door. It has none of what we normally expect a door to be, e.g., some sort of hinged, swinging slab of wood, cut into wall. The inner details are all wrong. Names and symbols, like analogies, are only partial truths; they work by taking many-leveled descriptions of different things and chopping off all of what seem, in the present context, to be their least essential details—that is the ones which matter least to our intended purposes. But still, what matters—when it comes to using such a thing—is that whatever symbol or icon, token or sign we choose should re-mind us of the use we seek—which, for that not-quite-door, should represent some way to go from one place to another. Who cares how it works, so long as it works! It does not even matter if that "door" does not open to anywhere: in TRUE NAMES the protagonists' bodies never move at all, but remain plugged-in to the network while programs change their representations of the simulated realities!

And strangely, this is also so inside the ordinary brain: it, too, lacks any real sense of where it is! To be sure, most modern, educated people know that thought proceeds inside the brain—but that is something no brain knows until it's told. W the help of education, a human brain has no idea that any such thing as a brain exists. To be sure, we tend to imagine our thoughts as in some vague place behind the face, because that's where so many sense organs are–yet even that impression is wrong: brain-centers for vision are far away, in the back of the head, where no naive brain would expect them to be.

An icon's job is not to represent the truth about how an object (or program) works. An icon's purpose is, instead, to represent how that thing can be used! And since the idea of a use is in the user's mind—and not inside the thing itself—the form and figure of the icon must be suited to the symbols that have accumulated in the user's own development. It has to be connected to whatever mental processes are already used for expressing the user's intentions.

This principle, of choosing symbols and icons which express the functions of things (or rather, their users' intended attitudes toward them) was already second nature to the designers of earliest fast-interaction computer systems, namely, the early computer games. In the 1970's the meaningful-icon idea was developed for personal computers by Alan Kay's research group at Xerox, but it was only in the early 1980's (through the work of Steve Jobs' development group at Apple Computer) that this concept entered the mainstream of the computer revolution.

Over the same period, there were also some less-publicized attempts to develop iconic ways to represent, not what the programs do, but how they work. This would be more useful for the different enterprise of making it easier for programmers to make new programs from old ones. Such attempts have been less successful, on the whole, perhaps because it is hard to decide how much to specify about the lower-level details of how the programs work. But such difficulties do not much obscure Vinge's vision, for he seems to regards present day forms of programming — with their stiff, formal, inexpressive languages—as but an early stage of how better programs will be made in the future.

I too am convinced that the days of programming as we know it are numbered, and that eventually we will construct large computer systems not by anything resembling today's meticulous but conceptually impoverished procedural specifications. Instead, we'll express our intentions about what should be done in terms of gestures and examples that will be better designed for expressing our wishes and convictions. Then these expressions will be submitted to immense, intelligent, intention-understanding programs that then will themselves construct the actual, new programs. We shall no longer need to understand the inner details of how those programs work; that job will be left to those new, great utility programs, which will perform the arduous tasks of applying the knowledge that we have embodied in them, once and for all, about the arts of lower-level programming. Once we learn better ways to tell computers what we want them to accomplish, we will be more able to return to our actual goals–of expressing our own wants and needs. *In the end, no user really cares about how a program works, but only about what it does— in the sense of the desired effects it has on things which the user cares about.*

In order for that to happen, though, we will have to invent and learn to use new technologies for "expressing intentions". To do this, we will have to break away from our old, though still evolving, programming languages, which are useful only for describing processes. But this brings with it some serious risks!

The first risk is that it is always dangerous to try to relieve ourselves of the responsibility of understanding exactly how our wishes will be realized. Whenever we leave the choice of means to any servants we may choose then the greater the range of possible methods we leave to those servants, the more we expose ourselves to accidents and incidents. When we delegate those

responsibilities, then we may not realize, before it is too late to turn back, that our goals have been misinterpreted, perhaps even maliciously. We see this in such classic tales of fate as *Faust*, the *Sorcerer's Apprentice*, or the *Monkey's Paw* by W.W. Jacobs.

A second risk is exposure to the consequences of self-deception. It is always tempting to say to oneself, when writing a program, or writing an essay, or, for that matter, doing almost anything, that *"I know what I would like to happen, but I can't quite express it clearly enough"*. However, that concept itself reflects a too- simplistic self-image, which portrays one's own self as existing, somewhere in the heart of one's mind (so to speak), in the form of a pure, uncomplicated entity which has well-defined wishes, intentions, and goals. This pre-Freudian image serves to excuse our frequent appearances of ambivalence; we convince ourselves that clarifying our intentions is merely a matter of straightening-out the input-output channels between our inner and outer selves. The trouble is, we simply aren't made that way. *Our goals themselves are ambiguous.*

The ultimate risk comes when our greedy, lazy, master-minds attempt to take that final step—of designing goal-achieving programs that are programmed to make themselves grow increasingly powerful, by self-evolving methods that augment and enhance their own capabilities. It will be tempting to do this, both to gain power and to decrease our own effort toward clarifying our own desires. If some genie offered you three wishes, would not your first one be, *"Tell me, please, what is it that I want the most!"* The problem is that, with such powerful machines, it would require but the slightest accident of careless design for them to place their goals ahead of ours–as were. The machine's goals may be allegedly benevolent, as with the robots of *With Folded Hands*, by Jack Williamson, whose explicit purpose was allegedly benevolent: to protect us from harming ourselves, or as with the robot in *Colossus*, by D.H.Jones, who itself decides, at whatever cost, to save us from an unsuspected enemy. In the case of Arthur C. Clarke's HAL, the machine decides that the mission we have assigned to it is one we cannot properly appreciate. And in Vernor Vinge's computer-game fantasy, True Names, the dreaded Mailman (who teletypes its messages because it cannot spare the time to don disguises of dissimulated flesh) evolves new ambitions or its own.

Would it be possible to duplicate the character of a human person as another Self inside a machine? Is anything like that conceivable? And if it were, then would those simulated computer-people be in any sense the same or genuine extensions of those real people? Or would they merely be new, artificial, person-things that resemble their originals only through some sort of structural coincidence? To answer that, we have to think more carefully about what people are—about the nature of our selves. We have to think more carefully about what an individuals is.

A simplistic way to think about this is to assume that inside every normal person's mind there is a certain portion, which we call the Self, that uses symbols and representations very much like the magical signs and symbols used by sorcerers to work their spells. For we already use such magical incantations, in much the same ways, to control those hosts of subsystems within ourselves. That surely is more or less how we do so many things we don't understand.

To begin with, we humans know less about the insides of our minds than we know about the outside world. Let me spell that out: compared to what we understand about how real objects work, we understand virtually nothing about what happens in the great computers inside our brains. Doesn't it seem strange that we can think, not knowing what it means to think? Isn't it bizarre that we can get ideas, yet not be able to explain what ideas are, or how they're found, or grown, or made? Isn't it strange how often we can better understand what our friends do than what we do ourselves?

Consider again, how, when you drive, you guide the immense momentum of a car, not knowing how its engine works, or how its steering wheel directs the vehicle toward left or right. Yet, when we come to think of it, it is the same with our own bodies; so far as conscious thought is concerned, the way you operate your mind is very similar: you set yourself in a certain goal-direction—as though you were turning a mental steering wheel to set a course for your thoughts to take. All you are aware of is some general intention—*"It's time to go: where is the door?"*—and all the rest takes care of itself. But did you ever consider the complicated processes involved in such an ordinary act as, when you walk, to change the direction you're going in? It is not just a matter of, say, taking a larger or smaller step on one side, the way one changes course when rowing a boat. If that were all you did, when walking, you would tip over and fall toward the outside of the turn.

Try this experiment: watch yourself carefully while turning—and you'll notice that before you start the turn, you tip yourself in advance; this makes you start to fall toward the inside of the turn; then, when you catch yourself on the next step, you end up moving in a different direction. When we examine that more closely, it all turns out to be dreadfully complicated: hundreds of interconnected muscles, bones, and joints are all controlled simultaneously by interacting programs that our locomotion-scientists still barely comprehend. Yet all that your conscious mind need do, or say, or think, is *"Go that way!"* So far as one can see, we guide the vast machines inside ourselves, not by using technical and insightful schemes based on knowing how the underlying mechanisms work, but by tokens, signs, and symbols which are entirely as fanciful as those of Vinge's sorcery. It's enough to make one wonder if it's fair for us to gain our ends by casting spells upon our helpless hordes of mental under-thralls.

Now take another mental step to see that, just as we walk without thinking, we also think without thinking! That is, in much the same way, we also exploit the agencies that carry out our mental work. Suppose you have a hard problem. You think about it for a while; then after a time you find a solution. Perhaps the answer comes to you suddenly; you get an idea and say, *"Aha, I've got it. I'll do such and such."* But then, were someone to ask how you did it, how you found the solution, you simply would not know how to reply. People usually are able to say only things like this:

"I suddenly realized..."

"I just got this idea..."

"It occurred to me that...

"It came to me..."

If people really knew how their minds work, we wouldn't so often act on motives which we don't suspect, nor would we have such varied theories in Psychology. Why, when we're asked how people come upon their good ideas, are we reduced to superficial reproductive metaphors, to talk about "conceiving" or "gestating", or even "giving birth" to thoughts? We even speak of "ruminating" or "digesting"—as though our minds were anywhere but in our heads. And, worst of all, we see ourselves as set adrift upon some chartless mental sea, with minds like floating nets which wait to catch whatever sudden thought-fish may get

trapped inside! If we could see inside our minds we'd surely say more useful things than *"Wait. I'm thinking."*

People frequently tell me that they're absolutely certain that no computer could ever be sentient, conscious, self-willed, or in any other way "aware" of itself. They're often shocked when I ask back what makes them sure that that they, themselves, possess these admirable qualities. The reply is that, if they're sure of anything at all, it is that *"I'm aware - hence I'm aware."* Yet, what do such convictions really mean? Since "Self-awareness" ought to be an awareness of what's going on within one's mind, no realist could maintain for long that people really have much insight, in the literal sense of seeing-in.

Isn't it remarkable how certainly we feel that we're self-aware—that we have such broad abilities to know what's happening inside ourselves? The evidence for that is weak, indeed. It is true that some people seem to have special excellences, which we sometimes call "insights", for assessing the attitudes and motivations of other people. And certain individuals even sometimes make good evaluations of themselves. But that doesn't justify our using names like *insight* or *self-awareness* for such abilities. Why not simply call them "person-sights" or "person-awarenesses?" Is there really good reason to suppose that skills like these are very different from the ways we learn the other kinds of things we learn? Instead of seeing them as "seeing-in," we could regard them as quite the opposite: just one more way of "figuring-out." Perhaps we learn about ourselves the same ways that we learn about un-self-ish things.

The fact is that the parts of ourselves which we call "self aware" comprise only a small portion of our mind. They work by building simulated worlds of their own—worlds which are greatly simplified, as compared with either the real world outside, or with the immense computer systems inside the brain: systems which no one can pretend, today, to understand. And our worlds of simulated awareness are worlds of simple magic, wherein each and every imagined object is invested with meanings and purposes. Consider how one can scarcely but see a hammer except as something to hammer with, or see a ball except as something to throw and catch. Why are we so constrained to perceive things, not as they are, but as they can be used? *Because the highest levels of our mind are goal-directed problem-solvers.* That is to say that the machines inside our heads evolved, originally, to meet various built-in or acquired needs such as comfort, nutrition, defense and reproduction. Later, in the last few million years, we evolved even more powerful sub-machines which, in ways we don't yet understand, seem to correlate and analyze to discover which kinds of actions cause which sorts of effects; in a word, to discover what we call knowledge. And though we often like to think that knowledge is abstract, and that our search for it is pure and good in itself–still, we ultimately use it for its ability to tell us what to do in order to gain whichever ends we seek (even when we conclude that in order to do that, we may first need to gain yet more and more knowledge).

Thus because, as we say, "knowledge is power", our knowledge itself becomes enmeshed in those webs of ways we reach our goals. And that's the key: it isn't any use for us to know, unless that knowledge helps to tell us what to do. This is so wrought into the conscious mind's machinery that it seems foolishness to say it: *No knowledge is of any use unless we have a use for it*

Now we come to the point of consciousness: that word refers to some parts of the mind most specialized for knowing how to use other systems. But these so-called 'conscious' ways to think do not know much about how those other systems actually work. Sometimes, of course, it pays to know such things: if you know how something works then you'll be better at repairing it when it breaks; furthermore, the better you understand a mechanism, the easier to find new ways to adapt it to other purposes.

Thus, a person who sustains an injured leg may begin, for the first time, consciously to make theories about how walking works: *"To turn to the left, I'll have to push myself that way"*—and then one can start to figure out,*with what could I push–against what*? Similarly, when we're forced to face an unusually hard problem, we sometimes become more reflective, and try to understand something of how the rest of the mind ordinarily solves problems; at such times one finds oneself saying such things as, *"Now I must get organized. Why can't I concentrate on the important questions and not get distracted by those other inessential details?"*

Paradoxically, it is often at those very moments—the times when our minds come closer than usual to comprehending how they themselves work, and we perhaps succeed in engaging what little knowledge we do have about our own mechanisms, so that we can alter or repair them—paradoxically, these are often just the times when, consciously, we think our mental processes are not working so well and, as we say, we feel "confused". Nonetheless, even these more "conscious" attempts at self-inspection still remain mostly confined to the pragmatic, magic world of symbol-signs. No human being seems ever to have succeeded in using self-analysis to discover much about how the programs underneath might really work.

I say again that we 'drive' ourselves–our cars, our bodies and our minds–in very much the self-same ways. The players of our computer-game machines control and guide what happens in their great machines: by using symbols, spells and images–as well as secret, private names. And we, ourselves—that is, the parts of us that we call "conscious"—do very much the same: in effect, we sit in front of mental computer-terminals, attempting to steer and guide the great unknown engines of the mind, not by understanding how those engines work, but just by selecting simple names from menu-lists of symbols which appear, from time to time, upon our mental screen-displays.

But really, when one thinks of it, it scarcely could be otherwise! Consider what would happen if our minds indeed could really see inside themselves. What could possibly be worse than to be presented with a clear view of the trillion-wire networks of our nerve-cell connections? Our scientists have peered at those structures for years with powerful microscopes, yet failed to come up with comprehensive theories of what those networks do and how.

What about the claims of mystical thinkers that there are other, better ways to see the mind. One way they recommend is learning how to train the conscious mind to stop its usual sorts of thoughts and then attempt (by holding still) to see and hear the fine details of mental life. Would that be any different–or any better–than seeing them through instruments? Perhaps—except that it doesn't face the fundamental problem of how to understand a complicated thing! For, if we suspend our usual ways of thinking, we'll be bereft of all the parts of mind already trained to interpret complicated phenomena. Anyway, even if one could observe and detect the signals that emerge from other, normally inaccessible portions of the mind, these would probably make no sense to the systems involved with consciousness. To see why not, let's return once more to understanding such simple things as how we walk.

Suppose that, when you walk about, you were indeed able to see and hear the signals in your spinal chord and lower brain. Would you be able to make any sense of them? Perhaps, but not easily. Indeed, it is easy to do such experiments, using simple

biofeedback devices to make those signals audible and visible; the result is that one may indeed more quickly learn to perform a new skill, such as better using an injured limb. However, just as before, this does not appear to work through gaining a 'conscious' understanding of how those circuits work; instead the experience is very much like business as usual; we gain control by acquiring just one more form of semi-conscious symbol-magic. Presumably what happens is that a new control system is assembled somewhere in the nervous system, and interfaced with superficial signals we can know about. However, biofeedback doe not appear to provide any different insights into how learning works than do our ordinary, built-in senses.

In any case, our locomotion-scientists have been tapping such signals for decades, using electronic instruments. Using that data, they, they have been able to develop various partial theories about the kinds of interactions and regulation-systems which are involved. However, these theories have not emerged from relaxed meditation about, or passive observation of those complicated biological signals; what little we have learned has come from deliberate and intense exploitation of the accumulated discoveries of three centuries of our scientists' and mathematicians' study of analytical mechanics and a century of newer theories about servo-control engineering. It is generally true in science that mere observational "insights" rarely leads to new understandings. One must first have some glimmerings of the form of some new theory, or of a novel method for describing processes: one needs a "new idea". Some other avenue must supply new magic tokens for us to use to represent the "causes" and the "purposes" of those phenomena.

Then from where do we get the new ideas we need? For any single individual, of course, most concepts come from the societies and cultures that one grows up in. As for the rest of our ideas, the ones we "get" all by ourselves, these, too, come from societies—but, now, the ones inside our individual minds. For, a human mind is not in any real sense a single entity, nor does a brain have a single, central way to work. Brain do not secrete thought the way livers secrete bile; a brain consists of a huge assembly different sorts of sub-machines parts which each do different kinds of jobs—each useful to some other parts. For example, we use distinct sections of the brain for hearing the sounds of words, as opposed to recognizing other kinds of natural sounds or musical pitches. There is even solid evidence that there is a special part of the brain which is specialized for seeing and recognizing faces, as opposed to visual perception of other, ordinary things. I suspect that there is, inside the cranium, perhaps as many as a hundred different kinds of computers, each with a somewhat different basic architecture; these have been accumulating over the past four hundred million years of our evolution. They are wired together into a great multi-resource network of specialists, in which each section knows how to call on certain other sections to get things done which serve their purposes. And each sub-system uses different styles of programming and different forms of representations; there is no standard language- code.

Accordingly, if one part of that Society of Mind were to inquire about another part, the two would most likely turn out to use substantially different languages and architectures. In such a case, if A were to ask B a question about how B works then how could B understand that question, and how could A understand the answer? Communication is often difficult enough between two different human tongues. But the signals used by the different portions of the human mind are even less likely to be even remotely as similar as two human dialects with sometimes-corresponding roots. More likely, they are simply too different to communicate at all—except through symbols which initiate their use.

Now, one might ask, *"Then, how do people doing different jobs communicate, when they have different backgrounds, thoughts, and purposes?"* The answer is that this problem is easier, because a person knows so much more than do the smaller fragment of that person's mind. And, besides, we all are raised in similar ways, and this provides a solid base of common knowledge. But, even so, we overestimate how well we actually communicate.

The many jobs that people do may seem different on the surface, but they are all very much the same, to the extent that they all have a common base in what we like to call "common sense"—that is, the knowledge shared by all of us. This means that we do not really need to tell each other as much as we suppose. Often, when we "explain" something, we scarcely explain anything new at all; instead, we merely show some examples of what we mean, and some non-examples; these indicate to the listener how to link up various structures already known. In short, we often just tell "which" instead of "what".

Consider how poorly people can communicate about so many seemingly simple things. We can't say how we balance on a bicycle, or how we tell a shadow from a real thing, or, even how one fetches a fact from one's memory. Again, one might complain, "*It isn't fair to complain about our inability to express things about things like seeing or balancing or remembering. Those are things we learned before we even learned to speak!*" But, though that criticism is fair in some respects, it also illustrates how hard communication must be for all the sub-parts of the mind that never learned to talk at all—and these are most of what we are. The idea of "meaning" itself is really a matter of size and scale: it only makes sense to ask what something means in a system which is large enough to have many meanings. In very small systems, the idea of something having a meaning becomes as vacuous as saying that a brick is a very small house.

Now it is easy enough to say that the mind is a society, but that idea by itself is useless unless we can say more about how it is organized. If all those specialized parts were equally competitive, there would be only anarchy, and the more we learned, the less we'd be able to do. So there must be some kind of administration, perhaps organized roughly in hierarchies, like the divisions and subdivisions of an industry or of a human political society. What would those levels do? In all the large societies we know which work efficiently, the lower levels exercise the more specialized working skills, while the higher levels are concerned with longer-range plans and goals. And this is another fundamental reason why it is so hard to translate between our conscious and unconscious thoughts!

Why is it so hard to translate between conscious and unconscious thoughts? Because their languages are so different. The kinds of terms and symbols we use on the conscious level are primarily for expressing our choices among and between the things we know how to do; this is how we form and express our various kinds of plans and goals. However, those resources we choose to use involve other sorts of mechanisms and processes, about which 'we' know very much less. So when our conscious probes try to descend into the myriads of smaller and smaller sub-machines which make the mind, they encounter alien representations, used for increasingly specialized purposes; that is, systems that use smaller and smaller inner "languages."

The trouble is, these tiny inner "languages" soon become incomprehensible, for a reason that is simple and inescapable. It is not the same as the familiar difficulty of translating between two different human languages; we understand the nature of *that* problem, which arises because human languages are so huge and rich that it is hard to narrow meanings down: we call that

"ambiguity". However, when we try to understand the tiny languages at the lower levels of the mind, we have the opposite problem—because *the smaller be two languages, the harder it will be to translate between them, not because there are too many meanings but too few. The fewer things two systems do, the less likely it is that something one of them can do will correspond to anything at all the other one can do.* Then, no translation is possible. This worse than mere ambiguity because even when a problem seems hopelessly complicated, there always can be hope. But, when a problem is hopelessly simple, there can't be any hope at all.

Now, finally, let's return to the question about how much a simulated life inside a world inside a machine could resemble our real life "out here". My answer, as you know by now, is that it could be very much the same—since we, ourselves, already exist as processes imprisoned in machines inside machines! Our mental worlds are already filled with wondrous, magical, symbol–signs, which add to every thing we "see" its 'meaning' and 'significance'. In fact, all educated people have already learned how different are our mental worlds than the "real worlds" that our scientists know.

Consider the table in your dining room; your conscious mind sees it as having familiar functions, forms, and purposes. A table is "a thing to put things on". However, our science tells us that this is only in the mind; the only thing that's "really there" is a society of countless molecules. That table seems to hold its shape only because some of those molecules are constrained to vibrate near one another, because of certain properties of force-fields that keep them from pursuing independent trajectories. Similarly, when you hear a spoken word, your mind attributes sense and meaning to that sound—whereas, in physics, the word is merely a fluctuating pressure on your ear, caused by the collisions of myriads of molecules of air—that is, of particles whose distances are so much less constrained.

And so—let's face it now, once and for all: each one of us already has experienced what it is like to be simulated by a computer!

"Ridiculous," most people say, at first: *"I certainly don't feel like a machine!"*

But what makes us so sure of that? How could one claim to know how something feels, until one has experienced it? Consider that either you are a machine or you're not. Then, if, as you say, you aren't a machine, then you are scarcely in any position of authority to say how it feels to be a machine.

"Very well, but, surely then, if I were a machine, then at least I would be in a position to know that!

Hah. That is a typically human, thoughtless presumption. It amounts to claiming that,*"I think, therefore I know how thinking works."* But as we've seen, there are so many levels of machinery between our conscious thoughts and how they're made that to say such a thing is as absurd as to say, "*I drive, therefore I know how engines work!"*

"Still, even if the brain is a kind of computer, you must admit that its scale is unimaginably large. A human brain contains many billions of brain cells—and, probably, each cell is extremely complicated by itself. Then, each cell is interlinked in complicated ways to thousands or millions of other cells. You can use the word machine for that but, surely, that makes little sense because no one ever could possibly build anything of that magnitude!"

Certainly, most persons regard it as belittling to be compared to a machine; it is like being treated as trivial. And, indeed, such a comparison is indeed insulting—so long as the name "machine" still carries the old meanings it had in times gone by. For thousands of years, we have used such words to arouse images of pulleys, levers, locomotives, typewriters, and simple other sorts of things; similarly, in modern times, the word "computer" has evoked thoughts about adding and subtracting digits, and storing them unchanged in tiny so-called "memories". Such terms no longer serve our purposes, because that tradition cannot describe how more complex machines could think like us. Accordingly, computer scientists have had to introduce a thousand new words for such ideas–and we've still barely started along that path.

As to the question of scale itself, those objections are out-of-date. They made sense in 1950, before any computer could store even a mere million bits. They still made some sense in the 1960s, when a million bits cost a million dollars. But, today, that same amount of memory costs but a hundred dollars (and our governments have even made the dollars smaller, too)—and there already exist computers with billions of bits. As I write this in the mid-1980s, some of my friends are building a computer that will be composed of more than a million smaller computers. When we finally discover how to make intelligent programs, the task of building the machines for them to inhabit will very likely be an already solved problem.

The only thing missing is most of the knowledge we'll need to make such machines intelligent. Indeed, as you might guess from all this, the focus of my own research in Artificial Intelligence is to find better ways to relate structures to functions through the use of symbols. When, if ever, will that get done? Never say "Never".