# 15 Paul J. Werbos

*Paul Werbos is Program Director of the Knowledge Modelling and Computational Intelligence program at the National Science Foundation, Arlington, Virginia. The best source for further reading is his article, "Optimization: A Foundation for Understanding Consciousness," in D. Levine and W. Elsberry, Editors,* Optimality in Biological and Artificial Networks?, *Lawrence Erlbaum Associates, 1996.*

## June 1993, Baltimore, Maryland

***ER:*** Paul, can you tell us something about your growing up and early childhood?

***PW:*** I was born on September 4, 1947, in the suburbs of Philadelphia. When I was very young, I was interested in mathematics. I remember getting a book when I was about eight, *All about the Stars,* from my parents. Then I went out and weeded lawns to earn money to buy books by Hoyle and Gamow, people like that.

The book that influenced me most was Fred Hoyle's paperback book. The title was *The Nature of the Universe* [Harper, 1951] and I remember reading it when I was eight, in 1955. I remember it because my family was very strongly Catholic, staunch Catholics. They originated, at least on my father's side, from the same neighborhood where von Neumann originated, around Timmesfarb, a German enclave in Romania. They were staunch Catholics, and my mother was a staunch Irish Catholic. I can remember when I read the first chapter of Hoyle's book, I was thinking to myself something really obnoxious like, "Oh, the glories of God's universe." And I got to the second chapter, and it got even better. Then I got to the last chapter, where it says there is this little planet full of these walking robots, and they had this incredibly arrogant idea that the universe was formed by a walking robot. They have these meaningless things called communism and capitalism that they shoot each other about, and they think they're important. And all of that is total illusion. All we have are the laws of physics.

My initial reaction to that was, "My God, this is what they've been telling me is evil. This is terrible. This is horrible."

And then I started thinking it over. How do I know? What is the basis for my knowledge? And the more I thought about it, I started saying, "Gee, my source of knowledge for Catholicism is what the nuns told me. This guy Hoyle probably knows more than those nuns do anyway." And so that was the end of Catholicism for me.

Independently, I became interested in mathematics. Obviously, mathematics fit into what Hoyle was talking about. I found in the attic one of my mother's old algebra books, and I snuck it away. When I told my elementary school people that I was already studying algebra, they laughed. But then somewhere around the fifth grade they decided, "Why don't we check? He does seem to know his multiplication tables pretty well." So they checked, and I did know algebra.

I finished the calculus course by the end of sixth grade. They sent me to the University of Pennsylvania to take the junior honors calculus course, complex variables and such, when I was in the seventh grade. My parents sent me to Lawrenceville in New Jersey so that I could go to Princeton. I had the equivalent of an undergraduate math major with a graduate course or two before I graduated at fourteen or fifteen from Lawrenceville.

During that time I took some summer courses in computing. This was one of the good things that NSF [National Science Foundation] did back then, but you couldn't easily measure results from it. It has since been discontinued because you couldn't measure it. I took one of those summer courses at Temple where I learned how to do computer programming. Then, the Moore School of Electrical Engineering at Penn. had an advanced version of the same thing, which I took. I guess I was fifteen at the time.

I worked for a summer before I went to Harvard. I worked at the University of Pennsylvania Hospital, Jefferson Hospital. We were supposed to be studying the circulatory system. The guy I was working for handed me D. O. Hebb's *The Organization of Behavior* [Wiley, 1949; a classic in the behavioral sciences]. I thought that was neat. I was supposed to give a talk back at the Moore School, to inspire the next generation of high school students. So I said, "How about I program this thing up [Hebbian learning], and I'll give a report on it."

We scheduled a date. But I looked at it, and I said, "Wait a minute, this is not clear." I tried different ways of looking at it, and I said, "This won't work."

It was obvious to me from a statistical point of view that Hebbian learning was going to be measuring correlation coefficients, and for multivariate problems it would not work. I never gave the talk. I said, "Now I'm going to figure out something with the same flavor that does work."

At that time, I guess I was sixteen. There were some other things that influenced me. I'd read the *Foundation Trilogy* [by Isaac Asimov, science fiction classics]. It had some ideas about organizations and thinking. I had forgotten that book, but when I reread it later, it was like a blueprint for my life. I didn't consciously remember it, but the impact of that book was incredible. I'd also read Feigenbaum and Feldman's *Computers and Thought* [McGraw-Hill, 1963]—a beautiful simple book. Ironically, Minsky's chapter was one of the things that got me most excited and turned me on. Minsky was one of my major influences. Well, Minsky and Hebb and Asimov and Freud.

I decided I wanted to do this. This will help us understand the human mind; it will help human beings understand themselves. Therefore, they will make better political decisions, and better political decisions are vital for the future of humanity. I remember explicitly thinking about how developing this mathematics could help us make better political decisions.

I got to Harvard and I was still interested, but there were no courses in it [brain theory]. There were math courses, but I already had them. There were computer courses, and I felt I already had them. Even the graduate courses didn't look as if they had anything that was really new. So what could I do? I took the one and only neurophysiology course they had. It was a course for premeds. You were supposed to spend a lot of time memorizing various hormones. I didn't do what I was supposed to do, so I got only a C+ in the course.

I spent a lot of the time reading Rosenblith's *Sensory Communication* [MIT Press, 1961], reading Morgan and Stellar [*Physiological Psychology*] and getting that vocabulary down straight and thinking about modeling. I spent a lot of my college career not doing what I was supposed to do but instead thinking about neural networks. I wound up majoring in economics.

You may ask, "What does economics have to do with it?" Well, by that time I'd decided reinforcement learning was the paradigm I wanted to pursue. I really felt reinforcement learning was the right way to think about it. I was still frustrated by the fact that I knew Hebb wouldn't work, and I wanted to find something that would work. Before you start fine-tuning it, you need a first-order thing that works. That's what we're still doing today. I feel as if I'm actually past that stage, but I feel the profession still has not reached the stage of having things that work.

I suspect if you take people in the audience here [the interview was held in Baltimore at an International Neural Network Society meeting], they're still doing supervised learning. They haven't really grasped reinforcement learning and what it means. I wanted to do reinforcement learning, and I knew it had something to do with distributed optimization because a neural net is a distributed system.

In economics, some people had spent years and years trying to figure out how to build an optimal, decentralized decision-making system. I was interested in that problem for its own sake, and it had larger political and social implications, which I still care about. I figured it's the same mathematics you need for neural nets. I wound up learning about linear programming, how that interfaces with economic systems, marginal thises and marginal thats, how you get nonlinear optimization out of a distributed system. My real goal was to build a mathematics to help you understand the human mind.

In senior year, one of my suitemates was Dan Levine. We all had single rooms on the gold coast in Adams House at Harvard, but he was the closest thing I had to a roommate. He and I spent hours and hours discussing neural nets. I'm the one who got him interested in them, and afterward he went to work for Grossberg in part because we had fun conversations. We used to debate optimization. He's going to be coming out with a new book [D. Levine and W. Elsberry, 1997. *Optimality in Biological and Artificial Networks."* Hillsdale, NJ: Erlbaum]. He says his real goal was to recreate our old discussions at Adams House about optimization and the human mind.

I wanted to publish some of this because even though I didn't have a working system, I felt as though this was a start toward a working system. Trying to publish the start of a new concept is very, very hard, especially if you're a mere student.

I was in London the year after Harvard, more for R and R than for anything else. I got a master's degree in London. That year I sent in a paper to *Cybernetica*, which was published in 1968. It was not a coherent thing, but it had the germs of all the later ideas. If somebody had ever had the mathematical ability and had pursued those ideas, a lot of this stuff would have been in the literature sooner than it was. Being published is not the whole thing. Maybe it was the wrong journal.

I talk in there about the concept of translating Freud into mathematics. This is what took me to backpropagation, so the basic ideas that took me to backpropagation were in this journal article in '68. I talked a lot about what was wrong with the existing [two state] McCulloch-Pitts neuron model, and how it was only "1" and "0." I wanted to make that neuron probabilistic. I was going to apply Freudian notions to an upper-level, associative drive reinforcement system. When I look at what Grossberg is talking about now, it's very, very similar because his gated dipole system basically is measuring a matrix. That's what I put in this paper. This paper showed how if you had that kind of matrix arrangement, you could derive a secondary reinforcement system that would allow you to do optimization over time. There's a link between the concepts in there and Grossberg's current concepts. In order to make it work from an engineering point of view, I later shifted to a different approach altogether, but it's possible that original approach could work. It's possible that what Grossberg and Levine are talking about could work. Nobody has yet applied the necessary engineering mathematics horsepower to find out if that kind of architecture could work because the engineers basically won't listen to what Grossberg writes.

***ER:*** Where did you go to school in London?

***PW:*** London School of Economics. To describe it as recreation may be a slight exaggeration, but Harvard was a problem-set kind of environment. There was a little bit of the treadmill, and, frankly, there was also a little bit of the monk, being an undergraduate at Harvard. If you wanted to spend half your time doing neural nets and still pass your courses, you wound up in a very monkish style of existence. Going to the London School of Economics was great for my sanity because they had a totally different style of education. It was a seminar-based system, not a lot of grades, and you learned a lot in class. We were doing international political systems. I wanted to understand these systems and see if that understanding could help do useful things in the world. That still is one of my major concerns.

We can complain about credit assignment in the neural network world, but in the political world it is ten thousand times as bad. [The "credit assignment" problem in artificial intelligence refers to the notorious difficulty of assigning credit or blame to the appropriate parts of a complex interacting system.] To get a good idea through the system, you have to abandon any pretense of credit, and then after you do that, you won't be able to continue on with the same endeavor. It's hard to apply systems theory to political systems—not for intellectual reasons, but for social organization reasons.

Then I went back to Harvard to get a Ph.D. in applied math. Having been fortified with humanity from London, I then descended into the bowels of the machine, stopping off at the RAND Corporation for the summer of 1968 along the way.

I was excited about working at RAND because I heard they were good in two things: U.S.-Soviet relations, and cybernetics and dynamic programming. I said, "My God, that's the right mix for me." Then I got there, and they told me I was going to work on the Vietnam War. I said, "Gee, it's not that I'm opposed to it; I have no moral inhibitions about it. It's just that you're going to get zero product from me because I don't know anything about Asia. I know lots about Europe, the Soviet Union, I know lots about mathematics, but if you give me work on the Vietnam War, it's just going to be crap. I'll do the best I can, but I want to warn you, the product isn't going to be good."

They interpreted that to mean that I was an evil war protestor. A guy named Ikle who later became director of ACDA [Arms Control and Disarmament Ageney] and then number three in the Pentagon personally threatened me that if I did not work on this Vietnam project and stop giving them caveats, I would be blackballed and would never have another job for the rest of my life.

At that point I said, "All right, all right, you want it, you get it."

Actually, I did learn a thing that summer that surprised me. It turned out one of their problems was coming up with a measure of success in the Vietnam War. A lot of people were using a measure called the body count. The body count was a total disaster. It came from a very highly classified paper, presumably declassified by now, which was total nonsense in any event. Niskanen, who later became head of the President's Council of Economic Advisors, made a very thorough, econometric study of the Vietnam War. Basically, he used factor analysis, which is a stupid method for this purpose. People then said, "Gee, enemy deaths are correlated with American deaths." They didn't say exactly that, but that's what drove the factor analysis. The prime factor therefore turned out to involve deaths of American soldiers. They didn't quite want to say it in these terms because basically it said the more soldiers you got killed, the better you were doing.

I looked at this, and I said, "OK, not only is this silly from a substantive point of view, but from a methodological point of view. Knowing something about systems design and control, I know why this is crazy. This is not a good measure of success."

So I immediately went into doing a dynamic system identification kind of thing, causal analysis, which was part of the adaptive critic design that I was into, that I even talked about some in *Cybernetica*—model-based optimization designs over time. I used all that to come up with an argument for better measures of success.

I looked for stable invariants, and they were things like Vietcong attacks on Americans. That was the best stable invariant underlying measure. I used that as a success measure and came to the conclusion that we should radically change our policy and do things like small-unit actions instead of these large sweeps.

They sent the paper to the Pentagon. The guy who was theoretically the principal investigator didn't discover it until very late in the game because of security and because he wasn't physically on site. So in September of that year, on my birthday, I wound up flying to the Pentagon, talking to the number three guy in the Pentagon. His name was something like Einthoven.

It was a strange conversation, because they walked in, and the number six man in the Pentagon said, "Uh, Mr. Einthoven, these guys have come up with an interesting result based on a statistical analysis of this data."

His first reponse was, "Bullshit." He just sat there. Meanwhile, everybody was shaking. He turned around and uttered two sentences. He said, "It's all the Marines, false correlation. It's I Corps."

Scurry around, scurry around, then number six guy said, "They did a separate split-sample study, and they excluded I Corps. They got the same results."

And Einthoven said, "Really? Well, we'd better look into it."

Of course, it was all classified, and I read it on the front page of the *New York Times* two weeks later, which immediately made me very cynical about American security. The minute it's useful, it's on the front page of the *New York Times*. That made me very cynical—that event plus a few other things that happened at the RAND Corporation. They were very diligent about checking your suitcases and having you show up to work at 9 point 00 point 00. When it came to major fundamental strategic matters, of course, that goes on the front page of the *Times*. What else do you do with important, critical things?

I went on to Harvard graduate school and descended into webs of problem sets. Mainly what I studied at the graduate school was mathematical physics and quantum theory. I minored in decision and control. I took Bryson and Ho's course and learned more about dynamic programming. After I had done the basic course work, I was torn because I wanted to do something very fundamental in science. There were two or three areas I was interested in. One had to do with the foundations of physics. But that was further out than backprop. Another area was models of intelligence. Another was models of motivation because you have to figure out where the reinforcement function comes from. I spent time thinking about all three areas.

I had passed all my course requirements. It seemed dumb to me that I had to sign up for four courses when I was supposed to be working on a thesis, but that was Harvard's rule. I decided to take a course where you get free computer time and do a computer project. Initially, I was going to do something on quantum physics. I learned something about partial differential equations, but not enough. I couldn't produce a really useful product at the end of $x$ number of months.

So I went back to the committee, and I said, "Gee I can't do that, but I have this little method for adapting multilayer perceptrons. It's really pretty trivial. It's just a by-product of this model of intelligence I developed. And I'd like to do it for my paper for this computer course."

They said, "Why don't you go talk about it to Larry Ho?"

I said, "Look, I've got a problem with this course. I can't solve the problem of reality in a course of six months, and so now I want to do a fallback. I've got this method for training multilayer perceptrons. I'm convinced it would work, and I know it's not a big thing, but at least I'll get credit for the course."

Ho's position was, "I understand you had this idea, and we were kind of openminded. But look, at this point you've worked in this course for three months, admittedly on something else. I'm sorry, you're just going to have to take an incomplete in the course."

And I said, "You mean I can't do it?"

"No, no, you'll have to take an incomplete because, basically, the first thing didn't work. We're very skeptical this new thing is going to work."

"But look, the mathematics is straightforward."

"Yeah, yeah, but you know, we're not convinced it's so straightforward. You got to prove some theorems first."

So they wouldn't let me do it. One of the reasons that is amusing to me is that there are now some people who are saying backprop was invented by Bryson and Ho. They don't realize it was the same Larry Ho, who was on my committee and who said this wasn't going to work. Ho was right to be skeptical because I was flying by intuition. If he couldn't reproduce my intuition in his head, it was entirely legitimate for him to be skeptical. I do think they should have given me permission, however.

By the time my orals came around, it was clear to me that the nature of reality is a hard problem, that I'd better work on that one later and finish my Ph.D. thesis on something small—something I can finish by the end of a few years, like a complete mathematical model of human intelligence.

So I defended before my thesis committee a mathematical model of intelligence and motivation. I said, "These are the things that I'm interested in. But I think intelligence is the one I'm going to do for the thesis."

I can remember those orals very well. It turned out the committee was more interested in motivation than they were in intelligence. I had a one-page prospectus that talked about each problem. Somewhere in that page, I think I made a statement that there might be parameters affecting utility functions in the brain, parameters that vary from person to person. You could actually get a significant amount of adaptation in ten generations' time. I was speculating that maybe the rise and fall of human civilizations, as per Toynbee and Spengler, might correlate with these kind of things. The political scientist on the committee, Karl Deutsch, raised his hand. I'd worked for him, by the way, in previous summers. He wrote *The Nerves of Government* [Free Press of Glencoe, 1963], arguing that the political system is a neural network system. He became president of the International Political Science Association. His book, *The Nerves of Government*, which compares governments to neural networks, is one of the classic, accepted, authoritative books in political science.

He raised his hand and he said, "Wait a minute, you can't get significant genetic change in ten generations. That cannot be a factor in the rise and fall of civilizations. That's crazy."

Next to him was a mathematical biologist by the name of Bossert, who was one of the world's authorities on population biology. He raised his hand and said, "What do you mean? In our experiments we get it in seven generations. This guy is understating it. Let me show you the experiments."

And Deutsch said, "What do you mean, it's common knowledge? All of our political theories are based on the assumption this cannot happen."

And Bossert said, "Well, it happens. Here's the data."

What happened was my oral defense became a discussion between the political science department and the applied mathematics and biology guys. I could scarcely get a word in edgewise. I passed the orals having said about two sentences and not having discussed models of intelligence.

I said, "OK, now I can do what I want to do because I passed with flying colors, even though I didn't say anything."

Then I got started. At some point, I had to write a prospectus on the model of intelligence. I did, and it was with an adaptive critic, and backpropagation was part of it. But the backpropagation was not used to adapt a supervised learning system; it was to translate Freud's ideas into mathematics, to implement a flow of what Freud called "psychic energy" through the system. I translated that into derivative equations, and I had an adaptive critic backpropagated to a critic, the whole thing, in '71 or '72. I actually mailed out copies of that paper to some people out at Stanford and certain people around Harvard.

The thesis committee said, "We were skeptical before, but this is just unacceptable. This is crazy, this is megalomaniac, this is nutzoid. So you have to do one of several things. You have to find a patron. You must find a patron anyway to get a Ph.D. That's the way Ph.D.s work."

I was under the illusion that getting a Ph.D. was your own creative piece of work. I'd read that. I believed democratic theory. I thought we were in a completely free country. I had lots of illusions back then.

"OK, I've gotta find a patron." They suggested that I go to MIT, and there were a few people the committee would accept. One of them was Grossberg. One of them was Minsky. One of them was Lettvin. I spoke to all three of them in a search for a patron. The committee didn't like the neural network area generally, but they said, "Look, you find a patron for your thesis, and we'll let you graduate in this area."

I went to speak to Steve Grossberg, who was an assistant professor at MIT. I remember this very wooden office. I walked in, handed him the papers, came back, and he said, "Well, you're going to have to sit down. Academia is a tough business, and you have to develop a tough stomach to survive in it. I'm sure you can pull through in the end, but you're going to have to do some adaptation. So I want you to sit down and hold your stomach, maybe have an antacid. The bottom line is, this stuff you've done, it's already been done before. Or else it's wrong. I'm not sure which of the two, but I know it's one of the two."

Well, if he'd said one or the other, I might have felt bad, but when he said, "I know it's one of the two, and I don't know which," I thought, "Gee, there might be a loophole in here somewhere."

Then he handed me some of his papers, and he said, "See, I have theorems to prove that it's already been done. So either you have replicated my work, or your work is wrong. I don't know which. But based on these theorems, I know that my work is the solution to those problems. So if you're willing to work within this approach, there might be something to do."

But I didn't hear that part, frankly because I was doing something else, and maybe there was a little pain in my stomach because clearly I had a problem. How was I ever going to graduate at this rate?

So he handed me his papers. I can honestly say, based on those papers—this was really early, like '71, '72—I know that he was talking about what we now call Hopfield nets because that's what was in these papers. He was having trouble getting it published. It may be that what happened with Grossberg is in part what happened with me—namely, we had the exact same idea in the exact same form, but people weren't willing to publish it. He had to dance it through and change it and modify it and screw it up before people would allow it to get through the system. And then after the screwed-up version got through, then people would allow the full form of it to come in from places that they trusted. We were not people they trusted, neither Steve nor I.

At any rate, what I heard from Steve was not encouraging. He did say, "You know, it might be nice if you had found an elegant thing like a LaGrangian formalism from which you rederive what I've got. That might be intellectually interesting if you found a way to rederive what I've got from a more general perspective."

I confess at that time I did not know LaGrangian mechanics. The funny thing was that the first course I'd ever taken in physics was quantum mechanics, and that was the only physics I knew.

I spoke to Minsky. I remember I had my Rosenblith, and I said, "You know, I've got a way now to adapt multilayer perceptrons, and the key is that they're not Heaviside functions; they are differentiable. And I know that action potentials, nerve spikes, are 1 or 0, as in McCulloch-Pitts neurons, but here in this book that I had for my first course in neurophysiology are some actual tracings. If you look at these tracings in Rosenblith, they show volleys of spikes, and volleys are the unit of analysis. This is an argument for treating this activity as differentiable, at least as piecewise linear. If you look at that, I can show you how to differentiate through it."

I went to Minsky for help, but Minsky would not offer help. Minsky basically said, "Look, everybody knows a neuron is a 1-0 spike generator. That is the official model from the biologists. Now, you and I are not biologists. If you and I come out and say the biologists are wrong, and this thing is not producing 1s and 0s, nobody is going to believe us. It's totally crazy. I can't get involved in anything like this."

He was probably right, I guess, but he was clearly very worried about his reputation and his credibility in his community.

Minsky also said, "You know, I used to believe in all this kind of stuff with reinforcement learning because I knew reinforcement learning would work. I knew how to implement it. I had a nice guy named Oliver Selfridge who came in and acted as my patron and gave me permission to do it. We coauthored a paper, but it was really my idea, and he was just acting as patron on the Jitters machine. I'll hand you the tech report, which we have deliberately never published."

It was his bad experience with the Jitters machine that turned him off on reinforcement learning and all the neural net ideas. It just didn't work. I later looked at that paper, and it was transparently obvious to me that what was wrong was that he didn't understand numerical analysis. He didn't understand the concept of numerical efficiency. We still have people in the learning business today who do not understand the concept of numerical or statistical efficiency. He had a system that was highly multivariate with a single reinforcement signal. The system can't learn efficiently with that. At any rate, he was totally turned off. That was the end of Minsky.

So I decided, "All right, now I'll try Lettvin." It was funny. I walked in, and he said, "Oh yeah, well, you're saying that there's motive and purpose in the human brain."

He said, "That's not a good way to look at brains. I've been telling people, 'You cannot take an anthropomorphic view of the human brain.' In fact, people have screwed up the frog because they're taking bachtriomorphic views of the frog. If you really want to understand the frog, you must learn to be objective and scientific. And besides which, even in physics, you know you can show the physical universe maximizes a utility function, but that doesn't prove it's an anthropomorphic entity. People have just got to get out of this whole style of thinking."

Besides, he said, "We will never understand the brain. It is too impossibly complex and ad hoc."

I recently ran across Lettvin, just this past year. He has mellowed in many ways. He recognizes that plasticity is what's really exciting about the brain. I'm an extremist about plasticity. He's even more extreme. I believe that within any layer of the nervous system, you have the same learning rule. Everything is produced by learning at the higher levels. The lower-level systems are so complex that we'll never get a total mathematical handle on them, but they're inherently boring anyway. At the higher levels, I would argue, there is an inherent modularity in the learning, which means that it is inherently understandable. A relatively simple set of learning rules explain the whole diversity of what we observe after the fact.

I believe it's just like the physical universe. You can't know everything in the physical universe, but you can understand the laws of dynamics, the laws of change. I would argue you can understand the laws of the learning that underlies the interesting stuff in intelligence. The cortex and the cerebellum, the olive—the fun places that provide higher intelligence. Lettvin is so extreme he even argues that one neuron can take over from another neuron, so it's really general and really modular. I don't think it's quite that general, but in terms of the learning mechanism, there may be histological development mechanisms that provide some additional flexibility.

I might add that when I was an undergraduate, I had a few conversations with McCulloch that influenced me a lot but had nothing to do with my Ph.D. thesis. Maybe the conversations with McCulloch changed my view of what neural nets were. I really got along with him. He was a really neat guy even though it was his model of the neuron I was challenging. Maybe if I'd talked to him about this model, I would have saved my career.

McCulloch changed my view of the human mind a little bit. I guess I should be totally honest, given the obscure nature of what we're about here. OK? This is a personal, not a scientific, sort of a thing. From age eight, I believed that when we understand the brain and we have the mathematics down pat, we'll get rid of a whole lot of mystical crap that has confused people and distorted their decisions and made them do bad things. As a result of some things that followed from conversations with Warren McCulloch, I eventually was convinced that maybe there are attributes of the human mind that we can't reduce down to neurons. I got so far off the deep end that I now go to Quaker meetings.

I think there is something out there beyond what our models of the brain are going to give us. It's ironic. The main thing Warren McCulloch did to my head was to send me into that orbit. But I still believe the mathematics is important because I believe that mathematics is a universal, just like the Pythogoreans used to say. I believe all forms of intelligence that we can possibly conceive of have to be governed by mathematics. Whether these forms are physical neural nets or not, I still think that mathematics is relevant. These days, if people ask me what my religion is, first I say Quaker Universalist, and if that doesn't work, I tell them to read Bernard Shaw's *Back to Methusalah*, which is the next best approximation. And then I actually discuss what I think is going on if they're intelligent enough to be interested in my idiosyncratic views of the mind. That's what Warren McCulloch did.

I didn't find a patron. Nobody would support this crazy stuff. It was very depressing. I tried to simplify it. I said, "Look, I'll pull out the backprop part and the multilayer perceptron part."

I wrote a paper that was just that—that was, I felt, childishly obvious. I didn't even use a sigmoid [non-linearity]. I used piecewise linear. I could really rationalize that to the point where it looked obvious. I handed that to my thesis committee. I had really worked hard to write it up. They said, "Look, this will work, but this is too trivial and simple to be worthy of a Harvard Ph.D. thesis." I might add, at that point they had discontinued support because they were not interested, so I had no money.

Approximately at the same time there were scandals about welfare fraud, about how students were getting food. They cut off all that kind of nonsense, so basically I had no money. NO money. Not even money to buy food.

A generous guy, who was sort of a Seventh Day Adventist and a Harvard Ph.D. candidate in ethnobotany, had a slum apartment that rented for about $40 a month in Roxbury in the ghetto. He let me share a room in his suite, in his suite with the plaster falling off, and didn't ask for rent in advance. I had no money at that time for food. There was a period of three months when I was living there in the slums. To conserve the little bit of money I had, I remember eating soybean soup and chicken neck soup. I remember getting the shakes from inadequate nutrition.

As for applying for jobs, this was in the days of the great aerospace layoffs. People were called overqualified, you know. I remember getting one very short-term job doing a computer program for an astrophysicist at MIT.

It was really terrible. I was a mile from the Harvard Medical School Library, where I would walk to keep my sanity every day. Past tons and tons of dog shit. They never cleaned the streets. And through the gangs. So I guess I was starving for my convictions.

Finally, they said, "Look, you know, we're not going to allow this." There was this meeting where we sat around the table. The chairman of the applied math department at that time was a numerical analyst, D. G. M. Anderson. He said, "We can't even allow you to stay as a student unless you do something. You've got to come up with a thesis, and it can't be in this area."

Karl Deutsch was the political scientist I had worked for, and he wanted to be helpful. He had a funny feeling something bad was happening here. Ideas like this shouldn't be totally destroyed. He didn't have the math, but he had a feeling something was going on here. He knew that I had done very good work for him in previous summers.

So Deutsch said, "You're saying we need an application to believe this stuff? I have an application that we could believe. I have a political-forecasting problem. I have this model, this theory of nationalism and social communications? What causes war and peace between nations? I have used up ten graduate students who've tried to implement this model on real-world data I've collected, and they've never been able to make it work. Now, do you think your model of intelligence could solve this problem and help us predict war and peace?"

I had actually seen that model and some of the earlier results. It was my conclusion at that time that some of the problems with that model were due to subtle stochastic effects, which I had worried about in a neural net context, but which also had conventional statistical aspects to them. I felt that the cutting-edge statistics that really relates to the neural nets—ARMA modeling—could do the job. My response to the mathematicians was to say what I just said. I said, "I can do it. It's not exactly a model of intelligence, but I believe I can handle this problem."

They nodded their heads, and they said, "All right. If you can use this new model of intelligence, and it actually works in predicting models of conflict—if you succeed, yeah, we'll give you the degree."

My impression was that half the guys felt, "Boy, this is a good way to get rid of him." And one or two felt, "Well, he'll do the statistical thing, but it will be legitimate, and it won't be any of this funny neural stuff."

So they said, "OK, go ahead."

The next thing, I went back to my Box and Jenkins and said, "Of course. The multivariate generalization is trivial; I'll go ahead and do it." And I went to the algorithm that was in the book by Box and Jenkins to implement it. I got computer time from the MIT Cambridge Project, which Deutsch was connected to. It was a joint MIT-Harvard project at that time. I said "I'll just code this up."

And then I looked at the algorithm, and suddenly, pain in my stomach because the cost of estimating multivariate ARMA processes, with the good, standard algorithm published by statisticians, increased like $n^6$. Even though $n$ was small, that would be enough to blow the computer budget. Suddenly, I felt very, very sick to my stomach. My God, I'm not going to graduate after all. I remember those days of sheer agony. It did hurt in my stomach a lot. And I'm not being figurative.

I remember pounding the walls with my mind and thinking, "Dammit, I can build a brain in order $n$. How come I can't do this faster than order $n^6$?" And I thought, "Wait a minute, wait a minute. Why can't I go back and use this little backpropagation algorithm and solve this statistics problem?"

Then I generalized backpropagation to handle time-varying processes—what people would now call recurrent or time-lag recurrent systems. I showed that I could use that to solve the statistical estimation problem within the allowed computer budget. So I went ahead.

The first application of backpropagation in the world in a generalized sense was a command that was put into the TSP [Time Series Processor] at MIT, available to the whole MIT community as part of their standard software. It was published as part of MIT's report to the DOD [the Department of Defense] and part of the DOD's report to the world. It was part of the computer manual, so the first publication of backpropagation was in a computer manual from MIT for a working command for people to use in statistical analysis. I was a second author of that manual. It was Brode, Werbos, and Dunn.

That manual went out. It said, "We're using this funny method," which I called dynamic feedback. One of the funny things was that we discovered that ARMA modeling was not the way to solve that statistical estimation problem. I'd gotten an idea for another way to solve the problem that did work, fortunately. The other method is something that the neural net community has not grabbed onto. I did a lot of nonneural tests of this alternate estimation method, which I called the pure robust method.

But, you know, you have to go one step at a time, and that's what I didn't understand. I got a lot of bad advice from people who said, "You know, you shouldn't have your next publications all be just your Ph.D. thesis. You've got to go on to do something new."

Well, that was bad advice because my Ph.D. thesis had enough for about five careers in it. What I needed to do was to publish one idea at a time, but it was so hard to get published that I tried to cram a lot into individual papers. I think that's one of the reasons they weren't widely recognized. Also, I felt very insecure about my access to journals because of the way people dumped on me in the past and the lack of encouragement, and that probably is why, like Grossberg, I have these early papers in the seventies that have twenty ideas—you know, two paragraphs on each, each of which will work—but that wasn't enough to really catch the eye of the community.

That's a digression. The new method worked. I did forecast nationalism and political assimilation. At that point, the Department of Defense became interested. In fact, when you can predict conflict twenty, fifty years in advance, it's amazing who can become interested. I had a prior track record from the Vietnam War that the community as a whole didn't know about, but they knew that "My God, this is the guy who told us how to get out of Vietnam," which was essentially true because that change in strategy led to a doubling in efficiency, which was what was behind our ability to remove troops from Vietnam to a great extent, although there are many other aspects to that story.

OK, so I had some brownie points at that time with the Department of Defense. I had two job offers after Harvard. I didn't know that the way the system works is your thesis advisor gets you a job. My advisor was Karl Deutsch, so my job had to deal with political science. There was no choice.

I went to the University of Maryland in what was supposed to be a public policy program. The provost had approved that. After I arrived, the approval had disappeared, so instead of being the quantitative guy in applied math and public policy, I instead found myself in a political science department proper, which was not a totally comfortable fit.

When I got there I also walked into being PI [Principal Investigator] on a major grant from DARPA [Defense Advanced Research Projects Agency]. It wasn't like I had filled out a grant application. It's like the DARPA guys came and spoke to me and said, "Now you are going to head a DARPA grant."

That's where the job came from, too, because the head of the department had been in charge of that office in DARPA—CTO, the Cybernetic Technology Office. It was a three-way grant. I was one-third PI on crisis management and forecasting. It was a three-year grant that bought off two-thirds of my time.

One third of my time was teaching. I would teach quantitative methods, like trying to teach backprop to graduate students in political science.

Two-thirds of my time was working for DARPA. They kept telling me, "Look, we don't want pure theory. You can spend some time on pure theory, maybe half your time. But you've got to do it as crisis warning. And we want a practical application."

I spent a lot of time working on adaptive critics. I already knew how to do heuristic dynamic programming [HDP]. Temporal difference methods are a special case of heuristic dynamic programming. But in the initial proposal to my thesis committee there was a statistical efficiency problem with having a scalar critic.

It was in this early period that I figured out how to get a multivariate vector critic, which I called dual heuristic programming, that solves the essential combinatorial problem. I published it in *The General Systems Yearbook*.

I also published the idea of heuristic dynamic programming. Barto and Sutton's TD [temporal difference] is a special case of HDP. It's fun, but it's not a model of the brain. After I'd done this theoretical work, DARPA was really pushing me for applications. "We need a real-world application of this stuff."

So I said, "OK, they want it real world. What's a real-world forecasting model?"

I found out that DARPA had spent a lot of effort building a worldwide conflict forecasting model for the Joint Chiefs of Staff that was used in the long-range strategy division of the Joint Chiefs. It was based on a worldwide data set that they'd spent millions of dollars on as the basis of global stratagy planning for the U.S. I said, "That sounds like a practical application. What we'll do is we'll take that model, which is based on something, which is sort of the equivalent of a TDNN [time-delay neural network], except classical. We will reestimate it using the pure robust method and the more advanced methods that I've derived since then in the same vein."

So I sent someone to get the database. First of all, the database was secret. Secondly, I was able to get it anyway. Third, it turns out that the data was grossly misleading. I think that is the right way of putting it. The fact of the matter is that it was largely interpolated data based on relatively unreliable sources. To make statistical causal inferences based upon this kind of data is highly improper in my view. I wound up sending a couple of graduate students to create a really good database of Latin America.

I said, "You want variance, high variance. Something hard to predict." I thought conflict in Latin America would be the most beautiful case. I figured there were enough cultural homogeneities that it would be a single stochastic process, but with lots and lots of variance in that process. So we got truly annual data going back, I don't know, twenty or thirty years for all the countries in Latin America and then reestimated the Joint Chief's model on it. It had an $r^2$ of about .0016 at forecasting conflict. By jiggling and jiggling we could raise it to about .05. It could predict GNP and economic things decently. Conflict prediction we couldn't improve, though; it was hopeless.

DARPA wasn't happy when I published that result. They wanted me to do something real and practical and useful for the United States. This was useful, but it was an exposé. They didn't like that. Therefore, the report, which included backpropagation in detail and other advanced methods, was not even entered into DOCSUB. Every time I tried to enter it into DOCSUB, somebody jiggled influence to say, "No, no, no, we can't publish this. This is too hot."

It was published in the *IEEE SMC [Systems, Man and Cybernetics] Transactions Journal* in '78 anyway because they couldn't block the journals, but it didn't include the appendices. So that paper in 1978 said, "We've got this great thing called dynamic feedback, which lets you estimate these things. It calculates derivatives in a single swoop, and you can use it for lots of things, like AI."

That was all in the paper, but the appendix on how to do it was not there because of page limits for a journal article. The fact that we could get better forecasts than conventional statistical methods was in the *IEEE SMC Transactions Journal* in '78.

At that point, DARPA was no longer happy. Things were getting uncomfortable at Maryland. Marxists wanted to take over the department and hated DOD people. I was not on the top of DARPA's list. I began to feel uncomfortable. Nobody was in my camp. I was just a lone, middle-of-the-roader surrounded by Marxists and military contractors and really traditional kind of political scientists. So, somebody offered me a chance to work for a year at the Census Bureau developing these ideas. I wouldn't commit myself to leave Maryland, but I spent a year at the Census Bureau.

We had reports for the Farmers' Home Administration that were just adaptive critics translated into policy language. The USDA was just about ready to put $1 million into using adaptive critics to allocate $20 billion a year of agriculture loans. They might have done it if I had accepted the job to stay on at the Census Bureau.

The Department of Energy [DOE] offered me a job. They said, "How would you like to be the person evaluating our global long-range forecasting models to tell us what is really true in the whole global energy policy area in order to advise the U.S. government?"

I said, "Gee, that sounds like an important job. Based on what is useful to the United States, maybe I ought to take that job." They wanted me to do the quantitative stuff too—maybe not as much as Farmers' Home Administration, but some. To this day, I don't know if I made a mistake. I accepted the job.

The guy who made me the offer is this wild guy, Charles Smith. He's a real personality. You may remember how *Alice in Wonderland* was based upon high-level people in the British establishment. There's a certain style that book tried to convey, and Charlie had much of that kind of style. At one point, he was a clubby at Princeton. He had good degrees. I think he had taught for Mosteller or Tukey [well-known statisticians]. He had a mathematical background, but he also had a unique personality as well—definitely the opposite of what you find in bureaucracy. Too far in the other direction, many people believe.

So Charlie hired me, and my job was to evaluate these models. I was sincerely concerned about energy. I had some credibility in that part of DOE because I wasn't just a guy trying to sell a methodology. They really, really wanted a sensitivity analysis of their very large long-range energy forecasting model, the official model used for long-range forecasting. They wanted to know how the inputs depend on the outputs.

They had a million dollar contract at Oak Ridge [National Laboratory] to study that model. They wanted me for several things. One, they wanted me to be a translator between engineering and economics. Two, they wanted a critique. They wanted exposés. They wanted me to rip apart the models of the Department of Energy in a very scientific, objective way that didn't look like I was trying to rip them apart, but was anyway. That's exactly what they wanted to hire me for, and I didn't really know that was the motive. These particular people didn't like modeling very much.

So at some point they wanted sensitivity analysis. And I said, "You know, I know a little bit about calculating derivatives."

Now the Oak Ridge model was not really a time series model, so they used their own sensitivity analysis methods, which they called adjoint methods. Historically, if you wanted to find roots outside of neural nets for backprop, you know, where I got it from was Freud, and maybe dynamic programming to a lesser extent. Another possible root, if you're looking for historical roots of things, would be the adjoint methods and people like Jacob Barhen. Jacob used to work at Oak Ridge.

Based on what Oak Ridge sent me, what they were doing was very different. A good way of describing it might be as follows. I had developed a technique to operate through time by way of arbitrary nonlinear sparse systems, dynamic systems. Oak Ridge had a technique for taking what we would call simultaneous recurrent nets and getting derivatives out of them, effectively but without really exploiting sparse structure.

It was about 1981 or '82 that I figured out how you combine time-lag recurrence and simultaneous recurrence and get derivatives efficiently out of both systems. I applied it to a natural gas model at the Department of Energy. In the course of applying it I did indeed discover dirt. They didn't want me to publish it because it was too politically sensitive. It was a real-world application, but the problem was that it was too real. At DOE, you know, you don't have First Amendment rights. That's one of the terrible things sombody's got to fix in this country. The reality of the First Amendment has deteriorated. Nobody's breaking the law, but the spirit of the First Amendment has decayed too far for science. At any rate, they finally gave me permission to publish it around '86 and '87. I sent it to the journal *Neural Nets*—that is, how you do simultaneous recurrent nets and time-lag recurrent nets together. Then the editorial process screwed around with it, made the paper perhaps worse, and it finally got published in '88, which makes me very sad because now I gotta worry about, "Well, gee, didn't Pineda do this in '88?"

So once again I had trouble working with journals. That's always been one of my problems. In fact, there is one case of a paper that I submitted to a journal where the reviews came back and said, "We can't publish this because it is a challenge to good people like Rumelhart."

But the paper wasn't negative. It just simply said, "I did this."

I've had a few other experiences that are much in the same spirit. We're probably better than a lot of other professions, but it's hard to stay objective when you have these experiences.

So that was in '81. I had an interesting opportunity then, when Charlie Smith was going to leave DOE. He was offered a job as director of the System Development Foundation. He told us, "I'm supposed to go out there and figure out how to combine weird things like brains, artificial intelligence, and math. Nobody's done it. Whoever gave us the money said, 'This is what we're going to fund.' You know, people who give money are sometimes a little crazy. So we're going to find legitimate things we can fund, but how can we fund such a weird thing?"

I went up to Charlie, and I said, "Charlie, let me give you a little briefing?"

I said, "Charlie, it's not as crazy as all that. I think there's a way to do it. And let me show you how. First of all, you've got this sensitivity analysis stuff. You know it works. You've seen it work. So why can't we apply it to a differentiable model of the neuron?"

I drew a little flow chart of where I thought the field was going and how this would fill a critical hole. And I said, "Charlie, this is what I think is do-able. You don't have to throw away the money. It can be done. And what's more, I'd like to participate in doing it."

To substantiate that I showed him a paper. Now, he was actually the reviewing authority for me. I couldn't publish anything without it going through the chain of command. You remember what I was saying about the First Amendment? In order for me to publish a nonenergy paper, it had to be reviewed only at the lower level, like a number two guy in the agency. If it was an energy paper, it had to go up to the highest level. Charlie Smith had signed off on a conference paper and on a tech report. Both of them, conference paper and tech report, described backpropagation in general terms for first and second derivatives, for sensitivity analysis, and for eigenvalues and applied it to energy models with substantial pages on applications to artificial intelligence and neural nets.

I said, "Charlie, you signed off on this. You understand the mathematics" —he certainly did—"so therefore you know that this could be applied. If you approved this paper, presumably the concept is reasonable."

He thought about it. He came back, and we had another conversation. I remember very vividly being in Charlie's office where he basically said, "Yeah, these are fine ideas and good directions to go, but if you want to do this stuff ... You are a civil servant. If you want to consider losing your tenure and your salary as a civil servant, and getting one third of the salary you have today for a job that ends in one year, with no security whatsoever beyond that, then I have some friends and I might be able to arrange something. But if you really want to work in this area long term, I mean, really, you are not the right person to do it because in a deal like this, we need to have the best people in the country with reputations. Otherwise, you're not going to be able to change the culture. I think you're not the right person to do this."

I don't know when he moved to California. I do know in the PDP books if you look in the acknowledgements section, there is acknowledgement of Charlie Smith and his critical role. Now, I obviously was not present in any conversations between Charlie Smith and anybody else. I do know I've heard people say, "Oh well, Charlie Smith didn't know any math." I know that much is not true. I mean you don't teach statistics at Harvard and Princeton without having at least some knowledge of math, and he did understand what this was about. But beyond that I don't know.

The other thing is that ideas can spread by nonverbal means. Even with the best intentioned people. When I was working at MIT and doing my thesis at Harvard, I remember attending a party once ...

You know, MIT guaranteed my survival. Once I started doing software for them, they discovered I was pretty good at it. Once I had put backpropagation into a TSP command, they offered me a full-time job at the Harvard Cambridge Project, so I did the Ph.D. thesis while having a full-time job.

While I was doing these things at MIT, I remember going to one party. I didn't go to a lot. And one of the people there said, "We have heard through the grapevine that somebody has developed this thing called continuous feedback that allows you to calculate all the derivatives in a single pass."

And I sat back and thought, "This says something about the grapevine." When you've got an idea that's hard for people to understand, it doesn't move fast. As soon as you've done what the whole world says you should do, which is distill the essence into a few understandable ideas, it spreads like wildfire. The people who have better access to the journals publish it before you do, which is a trap if you are a young person trying to develop good ideas.

After that party I decided that I would be a little closed lip for a while. After that, in 1981 in New York, I presented the conference paper that I mentioned, the one Charlie Smith signed off on. I presented it at the International Federation for Information Processing. IFIP was this gigantic conference. They seemed to know what I was saying. There was lots of applause. It seemed to be finally accepted. I mentioned we have applications, and now we can apply it to neural nets. I published the paper in their proceedings in '82, which became a Springer book. It's not like the usual conference papers, which never see the light of day. It was a real book and a real conference, and I had really talked at length about artificial neural nets and Grossberg and the limbic system and the brain. That's where I had the little diagrams with the circles of the multilayer perceptrons, only with reinforcement learning and backpropagation.

At that point, I thought, "Based on this publication, now my priority is assured, and I will relax a little bit. Now that I know that I've established that I've done it, now I will relax and let the grapevine hear about it. I will give the condensations, and I will send them out all over the world"—you know, even to people who have not heard of me because I knew how to jiggle the system to get the idea out.

I don't think it was a coincidence that shortly after I jiggled the system, all of a sudden things popped up. One place was with the System Development Foundation and California. Another place things popped up was MIT, and that was definitely one of the places I jiggled because I was running a contract at MIT at the National Center for Economic Research. I was contract manager there. I tried to get them to implement backprop. They said, "We'll send it around to the engineers at MIT and see if they are willing to implement this thing as a sensitivity analysis tool." They weren't willing to.

At any rate, the National Center for Economic Research is one of the places where my ideas popped up. I think that I did succeed at jiggling the system there. As contract manager, you know, you can do things like that. It is easier to create heresies from the top, sometimes, than it is from the bottom. One of the reasons I wanted to become a contracting officer was I'd seen how often our present system prevents new ideas from getting through. I wanted to be in a position to do the opposite, to take advantage of a position in the government to encourage what I felt was the best future direction. I think history has vindicated me. It was a legitimate direction to push. There's this incredible conservatism built into the system.

I had a disappointment with Charlie Smith. I stayed at the Department of Energy and got a chance to do a couple of papers. There was one on energy models and studies on the long-range economic modeling system at DOE and how you can use backpropagation to analyze convergence behavior of very large complex systems. That paper got good reviews in the operations research community. At least half the paper dealt with how you can implement backpropagation for complex energy models without using neural nets at all. That was '83. That's when I first had the idea of a dual subroutine, which I think is still crucial to the engineering implementation of these ideas. The guy who did the book that the paper appeared in was going to do something on factory automation, so I wrote a paper for that. He said, "Gee, it isn't real world enough for the factory; it's kind of general."

In that paper, I was putting the reinforcement learning in a new context, so he suggested I submit it to the *IEEE SMC Transactions Journal*. I sent the paper there, and it came out in January '87. That was the paper that Barto and Sutton read in January '87, and they suddenly realized there was a connection between what they had been doing and what I was doing. We got together in '87, very shortly after the paper was published.

I had a long talk with Rich Sutton and Oliver Selfridge at the GTE Research Labs in Waltham, Massachusetts. That conversation didn't work very well because Oliver Selfridge said things like, "I don't believe in any of this kind of crap. You know, I'm a good Anglican, and I believe in the soul."

I tried to reassure him by saying, "No, this is consistent, and I believe in the soul too. I just have a slightly different outlook on it."

That was not the way to handle Oliver. What I should have done was get down to the nitty gritty and show him how to design things. But I made the mistake of responding to the question, so we didn't get into the nitty gritty.

I think there were a lot of results from that discussion, at least with Rich Sutton. Rich was the one who mentioned Dave Parker to me, saying, "This guy has been doing backpropagation."

I got in touch with Dave Parker. Parker and Widrow were the people who invited me to the second ICNN [International Conference on Neural Networks] conference in San Diego in 1988. If it were not for Parker and Widrow, all these things—the advanced adaptive critics, the advanced estimation techniques, the generality of backprop theorems—would not be in the literature. We would be doing pure, simpleminded supervised learning until we quit and died from boredom. The next generation moving into neural control would not have happened if Parker and Widrow, for reasons of conscience, had not given me a chance to exist when a lot of people wanted to treat me as persona non grata.

To this day, there are people who are screwing up the discussion of recurrent nets and confusing people about recurrent nets because they don't want to cite my 1990 paper in the the *Proceedings of the IEEE*, where I describe backprop through time, very explicitly. I think that may be the best tutorial paper around on what backprop is—what backprop is through time and how to implement it. But they don't want to cite it, and so they'll cite the 1986 Rumelhart, Hinton, and Williams paper that deals with simultaneous recurrent nets. Then these poor guys will go out, and they think backprop through time has to do with simultaneous recurrent nets. They get mixed up between time-lag recurrent nets and simultaneous recurrent nets, which are really like night and day.

Widrow and Parker, on the other hand, were helpful. And when they gave me that opportunity, the very first thing I did in '88 was to say, "I thank you for the platform. Now let's talk about the real problem, which is not supervised learning. The real problem is intelligence. Intelligent systems, the mind, that's what I really want to do."

At that point, I was halfway well known, and NSF asked me to be a program director. I've been doing that ever since '88. I've been pushing the attempt to understand intelligence. When I took the job, 90 percent of my motive was to help us understand the human mind. My real ulterior motive was that I think the development of mental, spiritual potential is one of the imperatives in life. I think a better understanding of the mind in universal terms is crucial to that. I think that's the ultimate value of what we're doing. My goal in taking a job at NSF, even though it totally involves engineering applications, is to give me a chance to promote and develop the kind of mathematics that we need to begin to understand the human mind. I don't think a purely bottom-up approach by itself is going to do the job.

I've seen guys like Grossberg and Klopf do random searches through the space of models. It's clear that if you do that kind of random search, you get lost. There's just too much stuff, too many possibilities. You need to have a guideline, a magic trick that will lead you through the maze. Knowing what works and what doesn't is that kind of a guideline. The kind of mathematics we're developing does that. And so while the program is developing the mathematics, I hope eventually we'll come back and make the connection. It's going to be hard with the barriers that exist in our culture to make the connection from the real mathematics, the working mathematics, to the hard-core neurobiology and from there to psychology and from there to a greater appreciation of human nature.

These are hurdles we're going to have to go through on the way to changing the culture. It's clear our culture needs to be changed. It's clear that our understanding of the human mind is wrong in many fundamental ways. Mathematics can correct a lot of the basic fallacies.

Once I started on the job at NSF, it became apparent to me that the engineering applications were not only real, but some of them were truly important. At first I thought, "We'll make a billion dollars here and there for some company. There are a hundred technologies that are not being funded, each one of which could generate a billion dollars of product a year, no sweat. Neural nets are competitive, so they're one of those technologies."

But when I started learning about what some of the applications were, I began to realize, "Hey, these aren't just billion dollar applications. They could affect human history." I began to discover that the ability to achieve things like the human settlement of space required a technology that ultimately falls back to really tough nonlinear control problems.

I found out that greater efficiency in control can lead to major reductions in waste or pollution coming out of chemical plants. I found out that control or system integration is probably the biggest remaining challenge, along with manufacturing process control, in replacing the internal combustion engine with something clean, efficient, and sustainable.

After I learned all that, I began to shift my emphasis somewhat. At the present, I'd say that about half my motivation in the program is to push the kind of mathematical development that will help us understand the brain and the mind, and about half is to make a real and critical contribution to these major technological needs.

[Paul Werbos made the following additional comments in 1995.] Some of these areas have moved forward a whole lot, some seem to be just taking off now, others are still stuck. My greatest frustration is that there is such a huge amount of work still to be done.

The link to brain circuitry and new experiments has also grown stronger. We are talking seriously now about an emerging understanding of how intelligence works in the brain, an understanding that we are replicating in engineering. NSF has initiated two programs in engineering-neuroscience collaboration that open the door to funding this kind of development. NSF is also developing a still larger initiative that would strengthen the links to cognitive science and computer science as well. I have also initiated a small business program at NSF on fuel-cell electric cars, where neural networks have begun to contribute.

On the other hand, on the deep level of fundamental research and basic ideas, we still have lots of problems due to paradigm blinders and the walls between disciplines.

You would be shocked at what people can ignore even when it is staring at them in the face. We already have sketched out all the basic ideas we will need to build a truly brainlike, intelligent system, but filling in the holes will take a rare degree of creativity.

For myself, I've been drifting away from trying to fill in the big holes in the neural network area because I see some areas that are perhaps equally important that are much more in need of an early explorer. Issues involving basic physics hold most of my personal research attention. I have solved some of the problems that I worried about years ago, but a whole lot of follow-through is still needed.