Oral History of Bernard Widrow

Interviewed by:
Shayne Hodge

Recorded: April 15, 2013
Mountain View, California

CHM Reference number: X6808.2013

Shayne Hodge: Greetings, my name is Shayne Hodge. I am here with Dr. Bernard Widrow of Sanford University. Today is April 15th, 2013, otherwise known as tax day for everyone around here. We are doing this oral history at the Computer History Museum around 2:30 this afternoon and this is part of the oral histories being collected by the Semiconductors Special Interest Group of the Computer History Museum. So Dr. Widrow, welcome, it is good to have you here.

Bernard Widrow: Thank you.

Hodge: You had a long career in electrical engineering. You started off at your equivalent of a PhDs, Doctor of Science at MIT and then you wound up teaching at Stanford for about 50 years roughly. So why don't we start at the very beginning. As a child, did you want to go into engineering? Did you know what it was?

Widrow: Well, I will tell you how that happened. My father—well, first of all, I must tell you I came from a very small town in Connecticut called Norwich. This is not well known, this town. It is near New London which is much better known. That is a rival city of about equal size population. Population was about 35,000 people. So I grew up there, went to high school there and my father was the ice maker for this small town. He had an ice plant and the ice plant had electrical machinery. Whenever the electrician came to install some thing or fix some thing or do that I would be all over the guy asking him questions and he saw this little kid that was interested in what he was doing and asking questions. He was very patient in giving me answers and explaining things. So I started working with electrical things. The telephone company, local phone company, had a sort of a warehouse workshop, sort of focal place, not far from where I lived and they would be throwing away batteries and wires and little things like that and I began to connect dead batteries together and making a pretty good battery out of it and I would make little buzzers like a door bell buzzer. I made my own telephone, little things like that, that I was able to make from just pieces of stuff that I can find.

Hodge: In your father's ice making plant were there any electronics per say or was it just electrical?

Widrow: Well there was no electronics. Electronics in those days was radio and if you had a rad—we had a radio there that you could listen to but that was the extent of electronics in the ice plant. Everything is electric, electric motors, generators, synchronous motors, ammonia compressors, all sorts of things that I learned about. So that's how I became interested in such things and I think I was in the fifth grade and the teacher had a set of encyclopedias called the World Book, which still exists today. It was after class, I don't think I did anything bad. I think I just wanted to talk to her about something. So I was waiting for her after class. I was sitting there and there was this long row of encyclopedias. So I was just waiting. So I grabbed a volume at random and opened up a page at random and the volume I grabbed was the R volume. When I opened it up it was all about radio. So I started reading about radio. How does a radio tube work? So I read that and was just fascinated, totally captivated and I started to make radios when I was a little kid and I made radios and radios and radios and none of them worked. It turns out I didn't get the theory completely correct. There was a few missing concepts. The radios didn't work. Finally when I got the concept right I was able to make a radio that worked and I remember it was my sixteenth birthday and I had a work bench in the basement of our house, tubes all over the place, glowing tubes. My father came down and was able to listen with a pair of earphones and can hear speech from the radio and he looked at that and he scratched his head and he said, "I never taught him all that." <laughter> So I laughed and I was hooked.

Hodge: What year was your sixteenth?

Widrow: When a kid can do something that his father doesn't understand, this is good.

Hodge: <laughs> What year, sorry, was your sixteenth birthday, roughly?

Widrow: Well, I was born on December 24, 1929. So you have to add 16 years to that, if you can do that.

Hodge: So 1945.

Widrow: Something like that. But it was war time, Second World War. You couldn't get parts. You could not get parts for radios. Everything went to the war but there were a lot of old radios that weren't working anymore. I was able to pull parts from them and put together parts from this radio and that radio and put it together and make a radio that is working. So I was doing this for the neighbors. When I went off to college, I went to MIT. I had a very good friend in my freshman year, a man named Amar Bose, who is now very famous for his loudspeakers and Bose etcetera. He grew up in Philadelphia. We were both the same age within about a week. We were good friends and he told me that as he was a kid in Philadelphia fixing radios for all the neighbors just like I did. The big difference is that I did it for the love of it. When he fixed a radio for a neighbor he charged. <laughter> So he was the business man.

Hodge: What made you want to go to MIT?

Widrow: Well, I had a little chat with my father and at one point I—the man that I worshipped was the electrician who came to the ice plant to install things or fix things and talking about the future and I said to my father what I wanted to be was an electrician. My father says, "No, you're not going to be an electrician. You are going to be an electrical engineer." And I said, "What is that?" And he said, "I don't know what that is but <laughter> that's what they teach at MIT and you're going to go to MIT." <laughter> So small details like gaining admission. Well, one night, one evening, see during the war we were very short of help. My kid brother and I were working in the ice plant. We were kids. I mean these days OSHA would never have permitted anything like this. Working with heavy machinery and when you are making ice, you—every half hour you make what is called a pull. That means you pull ice out of the floor. You take the empty—you empty the cans out. The ice slides out. You fill it up with water and you put it back again. But there were six cans. Each one had 300 pounds of ice. So it is 1800 pounds of ice plus a couple of hundred pounds of steel in the cans. You got at least a ton of stuff you're fooling around with and these two kids are pushing this thing around, lifting it up with a crane. I don't know how we made it but it wasn't all that dangerous but it could have been. My father trusted us. So I was—one evening at the ice plant and it was a very big yard where all the ice trucks came in to get ice. A man came up; it must have been about six o'clock in the evening. He was driving a Cadillac. In those days that was quite unusual. A person driving a Cadillac was someone of great wealth, great distinction. It was a big black Cadillac. He was dressed in a suit, very nice and he saw me there. He called me over and we started talking. And we were just chatting and he said, "What do you plan to do in the future." I told him, "I want to be an electrical engineer." And he says, "Where would you like to go to study that." I said, "I'd like to go to MIT." And I told him it was really impossible. I am coming from a small town, no connections. We don't know anybody. I am coming from no where. I don't think I can ever get in. I'm not even going to bother to apply. He said to me, "You apply." He said, "If you don't apply, you'll never forgive yourself. You have to try. You can try and be rejected, that's okay. But if you never try, that's not okay." So I applied. I got in. So I came home from school—high school was close to home. I came home for lunch. My father was standing on the front porch waiting for me. Had a big, big envelope and he said, "The doctor at MIT wants to know about your state of health." And I said to my father, "Why the hell would he want to know that?" He says, "You dope. They accepted you." So off I went to MIT. I was scared to death.

Hodge: Do you remember who the man was in the Cadillac or why he was there?

Widrow: I never met him before, don't know what his name was, never saw him again. It was one of those funny things.

Hodge: So when you started at MIT, this would have been, I don't know, I guess '47 or '48 something like that?

Widrow: Forty-seven.

Hodge: So what was the state of electrical engineering education like at the time? What was sort of the distribution of courses you were taking?

Widrow: Most of the courses were about—for example, let's say we had a whole semester of DC Machinery, another semester of AC Machinery, a whole semester on transformers. How they work. How you design them. These are power transformers. We had machinery lab. We worked with pretty heavy duty, let's say 100 horsepower motors. And letting students play around with this stuff, the voltage was 550, three phase and you're plugging it in and out from a live panel. This would never be allowed today. So—but my—I had a roommate; a man named Dick Jenney, office mate, his father went to MIT. His father was MIT's patent attorney. So he went to MIT and then went to law school. When his father went to MIT they spent a whole year studying telephones, how to make a telephone, how a telephone works, all about telephone lines, every thing about telephones. So by the time I got to MIT there was nothing more about telephones but about vacuum tubes. I became a very good tube designer. We actually made our own vacuum tubes in tube lab. So it was quite different than what is today. But we were very busy. We had a lot to do. And you go to the MIT bookstore; you can buy a big banner that you could tack up on the wall that said, Tech is hell. It was. It was very hard. Long, long problem sets, lots of homework, work, work, work, work. When I was there, half of the class were kids like me right out of high school. The other half were guys that were in the war coming back under the GI bill. So they were typically about five years older than we were. So it was like two groups, the older guys and the kids. We were sort of the cut ups. They were very serious. They lost several years of their lives. They didn't lose their life but they were very, very serious about getting on with life. So it was really sort of a strange dichotomy and it was quite interesting, there was some amazing guys.

Hodge: As an undergraduate then were there any courses that you had signal processing transistors, computer architecture, anything along those lines?

Widrow: Are you kidding? <laughter> Computer architecture? <laughter>

Hodge: I had to ask.

Widrow: There were only six computers in the world at that time, approximately. MIT had one. Every computer was homemade, no two alike. There were no courses in it. When I was a graduate student, I did take a class in switching theory but it wasn't really computer architecture or computer electronics. It had to do with doing logic with relays which is what the telephone company was doing. The—there were—we had courses in electronics but it had to do with vacuum tube circuits, amplifiers. How you design them, this sort of thing, but until I was a senior, we—there was nothing about transistors. When I was a senior it turned out that MIT had two transistors and what we were doing was doing experiments reproducing some of the experiments that Shockley et al. had published in an article in the proceedings of the IEEE. So we were doing experiments, trying to reproduce what Shockley had done. So I worked with a transistor. I was able to make an amplifier from it and the amplifier was such that I couldn't even get a voltage gain of one. I couldn't even get a power gain of one. The input impedance was not infinite. It was fairly low in the kilohm range. The output impedance, the amplifier, was not low. It was also in the kilohm range. There wasn't much that the transistor could do, just allow, almost allow what was put in to go out and—but it was such that I knew that this was the future of electronics. It was just obvious to me. This tiny little thing, so small, compared to a tube and no filament, no standby power, just simple. And we had to be very careful doing these experiments because if you burned out the transistor you would have burned out half of MIT's supply. You would have burned out half of the transistors in all of New England. Where these transistors came from, came from Bell Labs. Somebody smuggled them out of Bell Labs and got them to MIT so we can do some special experiments. So that was my experience with transistors.

Hodge: Did you ever hear from Bell Labs about those?

Widrow: I never did. I was not responsible for smuggling them from Bell Labs.

Hodge: So that was in your senior year as a student there. What prompted you to stick around for graduate school?

Widrow: Well, at the graduation ceremony, there were people getting master's degrees. When I went to MIT I had no thought what so ever of graduate school. I didn't even know what that was. But I saw people getting master's degrees, begin to hear about it and also saw people getting doctoral degrees and the guys getting doctoral degrees had their academic costume with a hood, very colorful. I said, "Gee, I'd like to get one of those hoods and there is only one way to get it." <laughter> In any event, there were a couple of things about getting a mast—going for a master's degree. This was after, just after the Second World War. I was too young for the Second World War. When—the day that Pearl Harbor was bombed, I was 12 years old. So I totally avoided the Second World War. If it had gone on for another year or two I probably would have gotten drafted. I would have been fighting Japanese in the Pacific somewhere, like people older than me, the guys I knew. Everybody got drafted, everybody. My father could just barely keep the ice plant running. So the war was over. There was a huge economic down turn once the war was over. When the war was over there was no such thing as a concern about economy. Everybody was working, men, women, children. Everything went into the war effort. Once the war was over, 1945, '46, slowly all the war work and all the companies that were building stuff for the war, this all stopped and we had considerable unemployment. The year I graduated from MIT with a bachelors degree, 1951, twenty percent of the graduating class, my class, didn't get a single job offer. These are electrical engineers from MIT graduating with a degree in electrical engineering not getting a single job offer. So I never got involved in the job market but what I decided to do is see if I can get to go to graduate school. I was very, very lucky. I was able to get a research assistantship in a laboratory at MIT that was called the Digital Computer Laboratory. And there we were building a computer and I was assigned to the magnetic core memory group. So I was involved in the earliest development of magnetic core memory, which was crucial in making the computer practical.

Hodge: So what was the state of core memory when you joined the group?

Widrow: Well, there was no computer with a core memory. The computer at MIT used electrostatic storage tubes. They were highly unreliable and very, very costly, constantly having to be replaced. We had a whole group of people that had nothing to do except to manufacture electrostatic storage tubes.

Hodge: Is the principle of those like a capacitor?

Widrow: Yes, it is capacitor storage charge.

Hodge: This computer was it transistor based or tube based?

Widrow: My goodness, there were no transistors. <laughter> There were but…

Hodge: That's right you only had two.

Widrow: Yes, there were but the computer was called Whirlwind. Pieces of it today are in the Smithsonian. Whirlwind had something; I don't know five or ten thousand tubes, so racks and racks and racks of stuff. What that computer consisted of—the amount of logic, the amount of memory, the amount of computer, aside from the speed was—I think you can put 100,000 of them in your shirt pocket or in your wristwatch. So if you think at that time that we could have envisioned what happened with the computer. There was no way. We could never see it. So I would—going into this group, working on the computer, I would tell my father that this computer can add two—I would go home, sometimes I would go home on weekends. Where I lived was about 100 miles away from MIT. So sometimes on weekends I would go home and I was telling my father what it was that we were doing and I said this machine can add two numbers together. I am trying to remember the numbers but I think I said, "Can add two numbers together in ten millionths of a second." And he looked at me like I was crazy. He said, "Why on earth would anybody want to add numbers together so fast? <laughter> The way you add numbers together is you put them down on a piece of paper and you get your pencil and you add a column of numbers." So it was a total mystery to him why anybody would want to add numbers that fast.

Hodge: What was the application you were using the computer for at that time?

Widrow: The application for the Whirlwind computer was air defense. They had various excuses for building the computer. I think it was sponsored by the Navy, Office of Naval Research and I think the Air Force also was a sponsor. And the purpose was to take radar signals in real time, tracking aircraft, feed them into the computer, in real time and track them and be able to use outputs from the computer, in real time, to guide and intercept the aircraft, to go up and shoot down Soviet bombers.

Hodge: So was this your first introduction to signal processing or exposure to it?

Widrow: Basically, that is right. Unless you consider playing with radios, those were signals too. And a radio is a signal processor, isn't it?

Hodge: I haven't thought of it that way but yes I guess it would be.

Widrow: It really is.

Hodge: Well then I guess I should amend the question, your first experience with digital signal processors?

Widrow: Well, that was in graduate school. So what happened was that they were tracking aircraft, I was initially involved not with how they were applying the computer, but in how to make the computer. How to make parts of it and the part that I was involved in was the memory. And what we were building was random access memory, RAM.

Hodge: Did they have a prototype of it working before you were there or were you the…

Widrow: They had a little memory that can store four bits. And the first thing I did is design a memory plane that can store 256 bits. And I designed some of the electronics and this was all built up by technicians and then tested and we were able to, with crude programming, we were able to program it and put—store different patterns in the memory to make sure that regardless of the pattern, regardless of the nature of the bits that you were storing, that this memory had the capability of storing and retrieving. So the memory that I built, you could store things in about ten microseconds and it would take about five—to get information back from it would take about five microseconds. So that gives you an idea of the time frame.

Hodge: Was it read—in core memory was the read destructive of the data on that bit?

Widrow: It was destructive memory. What my master's thesis was a nondestructive memory for core memory. Nondestructive read out, that you could read it nondestructively. You see the way the core memory worked is you had—you store things by magnetizing these cores in one direction or the other. So you store a one or a zero. But in reading out from an individual core what the contents was, whether it was storing a one or a zero, what you always did was you wrote a zero into it. If the core already had a zero, you are not changing the state of the core. If the core had a one and you store a zero, you switch from a one to a zero. If there is a switching, there was a sensing wire, sensing wiring in the memory plane that would pick up a pulse because there was a rapid change in flux in one of the cores. So you get an output pulse or you wouldn't depending on whether the core that you selected was storing a one or a zero. But if you write a zero and the core was storing a one, you have to refresh, you have to rewrite the one. Otherwise you would have destroyed the information at that core and that was what you called destructive read out. So you had to rewrite it immediately. What I had done for my thesis, master's thesis, was to develop a method of nondestructive read out. So I did that. So in any event I was a graduate student when I did that and I was a part of this digital computer lab.

There was a class that I took from Professor William Linvill at MIT, a class called Sampled Data Systems, Sampled Data Systems. Today you call this digital data or in Europe they call it numerical data. It was taking a continuous signal and sampling it. And he taught a class in that. He, himself had done his doctoral thesis on that subject. Today you'd call that subject, digital signal processing and digital control systems. So it was a mixture of signal processing and control all in one class.

Hodge: Were you quantizing the signals too or were you just sampling them?

Widrow: The signals had to be quantized if they were digitized but the quantization was not treated. But what I did for a doctoral thesis was to develop a theory of quantization noise. That was my thesis.

Hodge: Did that correlate with your work in more of the application side of Whirlwind moving from the core to applications?

Widrow: There were no applications of digital signal processing.

Hodge: Well, let me rephrase that, of moving from working on the core memory to working on applications of that computer?

Widrow: The signal processing was totally independent of the core. In other words I sort of—what happened is that when I became a graduate student and took the class from Professor Linvill, I wanted to be his doctoral student. I didn't care what he was working on. I just wanted to be his student. I just liked the way he did things. I liked his thinking. I admired him tremendously and what he was working on was signals, algorithms. What I had been working on previously was hardware. So I switched from a hardware guy to an algorithm guy. I have been an algorithm guy ever since.

Hodge: Were you able to do any implementations of the algorithms or were they all sort of on paper analyses?

Widrow: This is all on paper. It is a funny thing, sometimes in academia you do things so many years ahead of the outside world. Long before any of this stuff has practical application, you are doing it. I mean to be 25 years ahead in a university, where you are doing basic research, is quite reasonable in an academic setting. You don't necessarily worry about immediate application. Today there's more immediacy than there was at that time. And so he was teaching a class and people were studying this stuff, learning about today you'd call fundamentals of digital signal processing and fundamentals of digital control systems. This was way before there were applications. I remember when I was still a graduate student or perhaps when I was a brand new faculty member, in another lab one of the guys was working on a project that would take an analog signal and convert it into digital. In other words, the world's first analog to digital converter. I remember when that was developed.

Hodge: Do you remember who that was?

Widrow: Yeah, Doug—oh, I'm having trouble remembering his last name.

Hodge: It might come back later in the interview.

Widrow: Yeah, I think if we looked in the history of digital analog converters, if we Googled it we'd probably find the rest of Doug's name.

Hodge: So you finished your PhD. You got the hood you wanted.

Widrow: Got the hood I wanted. See, the thing is when I did the master's thesis on nondestructive readout I needed a thesis advisor. So you had to do a masters thesis there. And I was taking the class from Professor Linvill. I asked him if he would be my thesis advisor and I showed him what I did. When he saw what I did, the trick I used to do a nondestructive readout, it wasn't his field at all but it was a form of signal processing. And when he saw that and I asked him to be my thesis advisor for the master's thesis he said, "Absolutely." So he became my masters thesis advisor. What I was really hoping to do was to go on for a doctorate, and he made that possible.

Hodge: So once you had your doctorate did you consider at that time perhaps going into industry or did you want to stay in academia?

Widrow: What happened was Professor Linvill, just as I was getting my doctorate, he was leaving MIT. He went to Washington, D.C. to work for IDA, the Institute for Defense Analysis. So he went there and I kept in touch with him, and I took over teaching his class. It's a class I myself took only two years before, but I already had some teaching experience. See, what happened when I was a first-year graduate student I took a class called Principles of Radar, and MIT was very big on radar because that was a big, big activity during the Second World War. And MIT was tremendously involved, MIT and Harvard, in the development of radar for the war. So I took a class from Professor Reintjes, Frank Reintjes, and he also was a man I admired greatly. I just liked the way he taught, liked the way he thought. And the year after I took the class, this was a graduate course, the year after I took it he needed somebody to teach the class and he asked me, "Could you teach it?" I said, "Sure." Never taught before in my life. Could I teach it? How do I know? But I said yes, so I did. That was one of the best classes, best bunch of students I've ever had. All the guys in the class were older people coming back from the war. They were all naval officers. They had all been to Annapolis [US Naval Academy]. They'd all been at the Naval Postgraduate

School in Monterey and they'd been out to sea, and they had operated radar sets and I'd never really had my hands on a radar set. And here I'm teaching Principles of Radar. These guys know all about radar. How am I going to teach them? They're going to be teaching me. It didn't work out that way because the course really had to do with the electronics of the radar, the electronics and the mechanical parts, all sorts of stuff about the radar. Things about that, what I taught in that class and what I learned from taking the class I'm still using today.

Hodge: So this was getting towards about the late fifties roughly at this point. One question for you. You've mentioned the guys from time to time. Were there any females in engineering at MIT?

Widrow: There were, there were. Let me describe that to you. I will describe it. In my freshman class there were 900 of us of which 11 were female. One was really good looking, and she lasted the freshman year and then got married to somebody and was gone. The others stayed and did fine, but it was really unusual. And if you needed young ladies for dating purposes what you had to do was go out to Wellesley, which was a good trip away, or Boston University or several of the schools around there. But basically MIT was a boy's school. Today it's completely different. I don't know. I think almost half of the undergraduates there are female.

Hodge: So were you formally offered a position as a faculty member at MIT or did you just sort of fall into it?

Widrow: Well, the answer is all of the above. <laughs> In other words, Linvill's class was very popular and there was no one there to teach it. So he recommended that I teach it and they agreed, and I became an assistant professor. So that's how I got into the teaching business.

Hodge: So how did you end up leaving MIT?

Widrow: Well, when I received the doctoral degree, this is a doctor of science, in those days the electrical engineers at MIT getting a doctoral degree, they got ScD instead of PhD, Doctor of Science instead of Doctor of Philosophy. The physicists, the mathematicians, they got PhDs. So PhD was something at the time was a little bit more prestigious than ScD. But they were all doctoral degrees. And today I think maybe ScD is a little more prestigious than PhD. PhDs are in everything all over the place. How many people do you know that are ScDs? Ha! You know me. Okay, so anyhow, I joined the faculty and began teaching that class and teaching other things but developing my own research. So I received the doctoral degree. Oh, I was going to tell you coming here the guys getting doctorates in electrical engineering, in my group there were 20 of us, 20 doctoral degrees in electrical engineering. Ten of the twenty were given the position of assistant professor, so what it was really was a training program for professors in electrical engineering, and there was no plan at MIT that those people were going to stay there, so they had to get a job somewhere else. Turns out that Professor Linvill had a brother, an identical twin, who was at Stanford, and he's the one that recruited me to come to Stanford, so that's how I got to Stanford.

Hodge: So did you visit Stanford before you accepted the job?

Widrow: Before I was offered the job I had made a trip to California to give a paper at a conference called Wescon. That's an IEEE conference. So I came and gave the paper. The paper was given in the Cow Palace. And to describe what the Cow Palace looked like at this huge IEEE convention it was amazing how they broke it up with canvas partitions. You can hear from one meeting to the other to the other, but it worked. And after the meeting I planned to spend a week in California. I'd never been outside of New England. So I got a Hertz rent-a-car and started driving at random. I didn't care where I went. I just wanted to see something of California. So I got in the car and started driving. And I knew that there were freeways and I didn't know what a freeway was. I'd been on the Merritt Parkway in Connecticut but nothing like a California freeway. The size and scale and the speed of the cars and the exits and the entrances and all this stuff. I didn't know anything about that, but somehow I got the car. I was driving in San Francisco and I saw signs that said 101. So I said, "Gee, that must be like Highway 1, which is on the east coast, north/south." Indeed, it is. So I got on Highway 101. Well, it turned out I didn't know that this thing goes two ways. There's a north and a south. Turns out I got on the direction going south. I didn't know that Hertz would give you a map. You could actually figure out where you're going. I had no map, nothing. Just got in the car and took off. This was, of course, before I was married, before I had children, you understand. Okay. So I started driving south. I began to go south and everything was fine, very interesting. I was just fascinated. Try to keep my eyes on the road but I'm looking all around and looking at all the stuff. I can't believe what I'm seeing. It's so different from the east coast. All of the sudden I began to see signs that said Palo Alto. I said, "Gee, that's where Stanford is." And I kept driving and I saw Palo Alto University Avenue, so I got off at University Avenue and I drove through Palo Alto and it looks almost the same driving through it today as it did way back then. This was in, we're talking about 1957-58, somewhere around in there. So I drove through Palo Alto, very, very pretty little town. Just nothing like that in Boston. And what a great place. And I drove down under El Camino onto Palm Drive right down into the Stanford campus. Then you could drive all around on the campus and you could see the whole thing. I said, "Boy, what a beautiful place that would be to spend a career." So I went back to MIT.

Hodge: So when you drove down 101 you said you were looking around. What was around 101 at that time?

Widrow: I don't know. Everything was different and fascinating. And after visiting Stanford, just looking at it, I spent overnight in East Palo Alto. At that time it was called Whiskey Gulch. They took it all out now. It's all gone. But I pulled up to a place that was a store that was selling prospecting supplies and right next to it was a Wells Fargo Bank. And I said, "My God, I'm in the Wild West. There's a Wells Fargo, prospecting supplies." <laughs> The 49ers, that wasn't a football team. It was the people that went out there digging for gold. The Wild West. And when I was driving down Palm Drive I saw palm trees. Palm trees. I never saw a palm tree in my life. I couldn't believe this thing. It was so beautiful, and the campus was so beautiful. So after a visit in Palo Alto and Stanford I took off on 101 and went south. I think—I don't know how far I went. Monterey, Carmel, further south. Went on Highway 1 a little bit here and there, and they had these roadside stands that were shaped like a big orange thing. And you'd come up and for ten cents you could get a tall, tall glass of freshly squeezed orange juice right from an orange tree out in the back of this thing. I mean, this was mind blowing. An orange tree. Who ever heard of such a thing? <laughs> We never knew where orange juice came from.

<laughter>

Hodge: Just came from the store?

Widrow: Came from a store.

<laughter>

Hodge: So that was sort of your self-recruitment out to California.

Widrow: Well, back before that when I first joined the MIT faculty, this was in 1956, the year of my doctorate, 1956, one of the guys at the lab said, "There's a conference going on at Dartmouth College right now, and it's an open conference. It's a- it's a workshop. Anybody can go. You can stay as long as you like. You can just listen or you can say something if you want, ask questions, visit with the people. It's on the subject of artificial intelligence." So I said to him, "What's that?" He said, "Some of these crazy guys think- think they're gonna build a machine that's like a human brain." I said, "Okay, I'll go. Let's go." So we took off and went to Dartmouth College, drive up to Hanover, New Hampshire. We stayed there for a week. All the original pioneers of the field of artificial intelligence were there, McCarthy and Minsky and several other people. Everybody was there except von Neumann. He wasn't there. I read a book about his life, and I now know why he wasn't there. At that time he was dying. He was very ill, and he was in his last days. But he did write some things about that subject, which I got chance to read.

Hodge: Was Wiener consider—

Widrow: Wiener? Wiener wasn't there. He was not into that. I can tell you Norbert Wiener stories that would knock you off that chair.

Hodge: We'll take one or two if you'd like.

Widrow: Okay. But that's another subject. You want to talk about Wiener. I was just going to tell you about that conference at Dartmouth.

Hodge: <inaudible>

Widrow: I came back after I had done my thesis on the theory of quantization noise, and I could tell you how academics works. You get on a subject, maybe your doctoral thesis, maybe something beyond that. And you work on that subject for the rest of your academic career. That's usually what people do. They may work on one or two subjects but not a heck of a lot. If you look at great artists, at great painters, you see, they get on a certain subject and they stay with that for a long, long time. Then they switch to something else. But if they're painting people's faces they do that. Or they do this, that or the other. They paint animals, objects, people, but that's what they do. And like Ansel Adams took pictures of Yosemite and the mountains, but he also, whether you know it or not, he took pictures of people, not very many, but he did. He took pictures of people, people's faces, in addition to his incredible landscapes. So what would've been natural for me would've been to spend my whole career working on the theory of quantization noise. I could've done that and made a whole life out of it, but I didn't. I wrote several IEEE papers on the theory of quantization noise and stopped working on quantization noise. I became so excited about the subject of artificial intelligence I could never get that monkey off my back. So there I was teaching a class in digital signal processing. It was Linvill's class that he called Sample Data Systems. In any event <clears throat> I spent about six months thinking about thinking, and I knew I really liked teaching. I liked the academics. I liked the academic life, the academic career. But I also knew that if I worked on something—I thought I'd be able to build or design a brain-like machine that would be working in about 25 years. I knew that if I worked on something and didn't have results for 25 years I'm not going to be able to stay in academia, not possible. You know, publish or perish or accomplish or down in smoke. What could I do in the short term never mind trying to build a brain? So I thought of a way to take a digital filter and allow it to learn, and what it was, was basically a Wiener filter that could adjust itself. And I had studied Wiener filter theory when I was a graduate student at MIT, so I knew all about that. So you develop a filter that will automatically adjust itself, self-optimize. It learns and it improves with experience, and I was able to build that and actually make it work by computer simulation.

Hodge: What did you build it on?

Widrow: This was done at MIT on an IBM 701 computer. It's a roomful of stuff. You'd call it a mainframe. Today you can put 100,000 of them into your wristwatch.

Hodge: Was Wiener at MIT when you were there?

Widrow: Oh, yes. He certainly was. He was the most famous professor when I was an undergraduate, when I was a graduate student and for three years when I was on the faculty at MIT.

Hodge: Okay. I understand his reputation was brilliant but somewhat eccentric.

Widrow: To say the least. All of the above. You're absolutely right. He was brilliant, but he was like a teddy bear. He had like a child-like personality. So we'd see him walking down the long corridors at MIT smoking a cigar at angle theta, typically around 45 degrees, and he'd be puffing away on the cigar and his head is encased in a cloud of smoke. And there he is walking down this long, long hallway. At the end of the hallway there are steps and they have steel edges so they're concrete and steel, very unforgiving. So Professor Wiener, he had a huge body. He was about, I don't know, maybe five foot five inches tall, but I'm sure he weighed about 300 pounds. He was an enormously heavy guy. He could never see his feet. When he walked down the corridor he's looking up. See, he's puffing on the cigar but he's looking up, and all he could see is the ceiling and the walls, the boundary between the walls and the ceiling, and that's how he's able to guide himself down the hallway. So people would see him coming and just get out of his way because he doesn't know they're there. He's just puffing away. He's deriving equations, you understand? So there he is. He's coming to the end of the hallway and I know this is the end of the hallway. He's not looking. He's puffing on the cigar. He's deriving equations, so you could say. So if you interrupt him and say, "Professor Wiener, you're at—be careful. There are steps." But if you do that, you interrupt his thought, you might set science behind by ten years. On the other hand if you don't tell him about the stairs he might kill himself. So what do you do? So you just stand there and pray.

<laughter>

And he's done this year after year and never fell down the steps, somehow. So another story. You want two Norbert Wiener stories. I'll tell you another Norbert Wiener story. This one's really cute. He loved it when people recognized him. He wanted to be famous. He loved being famous. He loved it. It was obvious, and you know, you'd see him everywhere. He's all around the place. He's wandering all over. He's not sitting in his office. He's puffing the cigar and walking the big hallways. So one day, it was Sunday and they didn't serve food in the dormitory, so I was with two of my closest friends and the three of us were going out. I had a little car, a little Plymouth, and we were going to go to get lunch, brunch. It was brunch time. So we got in the car and drove up on Memorial Drive up to a restaurant called the Smith House. So going to the Smith House was a real treat compared to dormitory food. So I parked the car, the three of us got out, we went into the Smith House. We went into the entrance and it had a lunch counter and it had a set of booths, and the booths were alongside a wall with windows that look out over the Charles River, so it was quite nice to get a booth there. So we came in and at the lunch counter who do we see sitting on a stool having brunch? Norbert Wiener. He's an enormous man on this little tiny stool. Yet he's sitting there and he's having a meal. So we went in, got a booth, had brunch, paid the bill, got up and we were going out. And we noticed when we went out by the lunch counter that Wiener was gone. So I had the keys to the car. I was going outside to the parking lot. We're going along and we see this old battered Chevrolet, old beat up car. And who's in the car? Norbert Wiener. So I walk by and he rolled down the window and he said, "I say, I seem to be having trouble starting my car. I wonder if you could give me a push." And I said, "Certainly, Professor Wiener. I'd be very happy to give you a push." As soon as I mentioned his name his face lit up he was so happy. I think he must've assumed looking at us that we were three MIT boys at the Smith House. Well, I have to tell you before I continue with the story a little bit of background. He wrote a book called “Cybernetics,” and now we got cyber-this and cyber-that and blah, blah, blah, blah, blah. Nobody knows where it came from and I can tell you. Cyber is from Norbert Wiener, his book on cybernetics. What the book was about was about feedback and feedback control in the body of humans and animals, how nature does control. It’s man, machine control. It’s a whole hodge-podge of ideas that he had. So he knows all about control and all about feedback in animals, et cetera, et cetera. Okay. Now I’ll continue with the story. So he said, “Can you give me a push?” And I said, “Professor Wiener, I’d be glad to- to do that.” So I got in my car and drove around, got behind his car. In those days you could start a car with a push. See, these were not automatics. Everybody had a stick shift. So I got behind his car. I put my car in first gear so I’ve got plenty of get-up-and-go. And I got behind and I started letting the clutch out and giving gas so I don’t stall my engine, and his car is not moving. So I said, “I know what he’s doing. He’s got his clutch out and he’s got it in first gear.” I said to the guys in the car, I said, “Should I go and try to explain to him how you- you should handle a car getting pushed, explain how you do a push?” They said, “He’d never understand.” So I said, “He’d never understand,” so I said, “Look, I’m gonna tough it out. I hope I don’t burn out my clutch.” So I give it the gas and slowly let out the clutch and then gradually inch by inch his car starts to move, and then it starts to move a little bit, little bit, little bit and all of the sudden he takes off like a rocket. So I know exactly what he did. He had it in first gear. Should get a push in second gear or third gear and get the car rolling and then let the clutch out once your car is rolling, but you don’t start from a dead stop with the clutch out in gear. Try to force who’s pushing you to get—well, this is another story. Anyhow, so he’s in first gear and he’s got his foot on the gas pedal and he takes off like a rocket. And of course, being a heavy guy the car shoots out from under him and it pulls him back. And I guess his weight came off the gas pedal, and when you’re in first gear and you take your foot off the gas pedal the car practically comes to a stop. So his body goes forward, he gets his foot again now on the gas pedal and he goes Vroom! And then again pulls him away from the gas pedal, so he’s going Vroom! Vroom! Vroom! Vroom! Like that. Meanwhile he’s waving his hand to thank me and he’s blowing his horn to thank me, thank me, thank me and he’s going Vroom! Vroom! Vroom! And there he goes out of the parking lot right onto Memorial Drive. Oh, my God! There’s two-way traffic, two lanes one way, two way the other way and they’re going fast. I said, “My God!” I said to the guys in the car, I said, “We killed him.” He goes out and the horns start blowing and the brakes are squealing and screeching, and he’s still going Vroom! Vroom! And he’s waving his hand thanking me. <sighs> So he survived.

Hodge: <laughs> He modeled an unstable system.

Widrow: So in other words he’s supposed to understand about feedback. That’s what he was having. He was oscillating with feedback. Anyway. So I tell you, he was brilliant, but about practical things in the real world, forget it.

Hodge: I’m going to skip forward just a little bit here. I’ve got your CV in front of me, which is rather long. You’ve published quite a bit during your career, but one of the things that stands out to me is this paper suddenly, there’s nothing like it before. It’s on here. “B. Widrow and M.E. Hoff, ‘“Associative storage and retrieval of digital information in networks of adaptive neurons," 1962. So, I take it at this point you're at Stanford, and you have picked up a graduate student named Ted Hoff.

Widrow: Yes. I'll tell you how that happened. When I first came—first began teaching at Stanford, I was in my office, received a phone call from John Linvill, the twin brother of my thesis advisor. And John was already at Stanford, came before I did. John Linvill was recruited to Stanford by Shockley—well Shockley suggested this to the powers that be at Stanford, to get John Linvill. So, they stole him away from Bell Labs to start a semiconductor activity at Stanford in electrical engineering. That's what brought John Linvill here. Anyhow, John Linvill was a—had one of his advisees was a man named Ted Hoff. And the way John described it to me is that he explained to Ted all the things he was doing in his research. But it doesn't seem like any of it really attracted his attention. And he said would you be willing to talk with him to see if any of the work you're doing could be of interest to him. And if so, then he could become your doctoral student. So, I said okay I'd be happy to speak with him. He said he thinks—he's very quiet. He's a very, very bright guy. And he said I can't quite figure out what he's really interested in to do for research. So, I said sure send him along. I'll be happy to talk with him. So, the first session I had with Ted, I was up at the blackboard explaining ideas of an adaptive filter, a filter that can learn all by itself. And in explaining that to Ted, the course of the discussion, out came the idea of an algorithm called—today called the LMS algorithm. It stands for least mean square. That algorithm is used—is the most widely used learning algorithm on the planet today.

Hodge: Now, this came out of a discussion before he was even your student—actually your student?

Widrow: Well, I don't know what happened. The two of us just became attached together. And the first session we had, out came that algorithm. And within about a half hour after I wrote it on the blackboard, he had it programmed on an analog computer that was directly across the hall from my office. We had it up and working. Then we started to think of how instead of using that big analog computer that takes up a whole room, how we can build something small I can just have on my desk in my office. So, we designed and then realized it was Friday afternoon. And the stock room was already closed. So, we couldn't get any electronic parts. The next day was Saturday. So, it wasn't going to be open. So, the next day we went downtown in Palo Alto to an electronic shop called—what the heck is the name of it? I keep thinking of Fry's, but it's not Fry's. Fry's didn't exist at that time. It was Zack, Z-A-C-K, Zack Electronics. They still exist. I think they moved from Palo Alto to Sunnyvale years ago. We bought all the parts. And that weekend we put together the first adaptive machine. We made a learning machine that we can teach to recognize patterns. And that learning algorithm is used today in every modem in the world.

Hodge: Now—

Widrow: And I can tell you that without that, we would have had digital communication on telephone lines at the rate of about 300 bits per second. And that would be it. But now, because of that algorithm, we have what's called adaptive equalization and adaptive echo canceling, canceling echoes in digital transmission. Without that capability, I don't think you'd have an Internet.

Hodge: The device you built, was that Adaline [sometimes written as ADALINE]? Or was that a precursor?

Widrow: That's Adaline.

Hodge: Okay.

Widrow: And what I'm describing is the first Adaline.

Hodge: Could you describe the device briefly?

Widrow: Yeah. It was—it's what I call a linear combiner. So, you have many inputs and a single output. It's the equivalent of a single neuron and the synapses leading into it. So, we have variable weights, variable coefficients. And these—they're like volume controls, they're potentiometers. And they act like as we understand how synapses in living animals work. It's an electrochemical device that is conductive. And the conductance of each synapse varies with experience.

Hodge: Was the biology of that well understood at that point?

Widrow: Well, not like today. There were questions whether the synaptic—whether the synapses, whether there were changes in the conductance of the synapse. Now, we know there are. So, it was controversial. But it was inspiring to us. And so, we developed this adaptive circuit. And we were able to demonstrate learning by feeding in input patterns and teaching it to respond and classify those input patterns. So, all input patterns would be classified into two classes, Class A and Class B. And once you've trained it, then you can put those patterns in. And it would give you the output responses that are representative of Class A or Class B, whatever the case may be. And there were no applications for it. Pattern recognition was nothing that really existed.

Hodge: I could be totally off base here, but this sounds a little bit like finite impulse response digital filter in terms of for the sum of weighted inputs. Is there a similarity?

Widrow: You're exactly right. I'll explain that. There were two applications for this that we developed over time. And one was for pattern recognition. The other was for adaptive filtering. The application in modems, that's adaptive filtering. So, that's a tap delay line with variable weights—variable tap weights.

And by varying that, you're varying the impulse response of the filter and allowing it to self optimize. It learns to improve itself with experience.

Hodge: It learns what filter coefficients are best to—?

Widrow: That's right.

Hodge: Okay.

Widrow: Yes.

Hodge: And so, as you said, that's obviously all around us and particularly in electronic communications. When was the first time you went—as I understand, Adaline was sort of a special purpose analog computer in a sense.

Widrow: Yeah. I wouldn't call it an analog computer because an analog—when you say analog computer, you're talking about a certain structure. We were able to do the first Adaline on an analog computer. But Adaline itself is not an analog computer.

Hodge: Okay.

Widrow: I had been working on, while I was still at MIT, I'd been working on adaptive filters. But the method of adapting—these methods are based on the idea of using a gradient. We were trying to minimize an error function. You're trying to minimize error. And the function of the error that we were working with was mean square error. This was the error criterion that Norbert Weiner used. I think the idea of minimizing mean square error came from Gauss. And it turns out that I am related to Gauss in this way, academically. There's a website that you can look up, put your name on there, and it's there. Then you see who your thesis advisor was and see who his thesis advisor and who's his thesis advisor. If you go back, back, back, back many, many generations, my line goes right back to Gauss. And beyond Gauss, it goes further back to Liebnitz.

Hodge: Impressive lineage.

Widrow: So, that's my lineage.

Hodge: So, you implemented Adaline. At some point in this process, you actually wound up creating an electronic component, a variable resistor with memory.

Widrow: That's right. That's called the memistor. It's a resistor with memory.

Hodge: That's not to be confused with the memristor, correct?

Widrow: It's a completely different device. The memistor was about ten or twelve years before the memristor. The memristor is sort of an awkward name. Memistor is a very smooth nice name. I think he had to call it memristor because memistor was already spoken for. But I never met the guy and never talked to him. I'm just guessing. Anyhow, the first Adaline we made didn't have memistor. It had potentiometers that were mechanically variable. And so, I had been working, long before I came to Stanford, on adaptive neurons and adaptive filters, variable weights. But the method of the algorithm was quite different. Once we got LMS, much, much simpler, much more efficient. And that's the algorithm that today is used. The methods I had before were very crude, simple methods of finding the gradient by taking each component and rocking it forward and backward to see whether performance got better or worse and always changing the coefficients in the direction to make them better. With LMS, you didn't rock them back and forth at all. You got a formulation for how to change the weights to make it better. And you did it all at once, bang. And that we couldn't do before LMS. We had to do each weight one at a time.

Hodge: At what point did you first try to implement one of these, I'm going to recall the learning machine, the digital computer instead of a—?

Widrow: Oh, that came quite soon after, maybe six months. The first things we did were we're building circuits, building circuits to do it and, at the beginning, mechanical potentiometers. Then thinking about living neurons and synapses, the synapse in the living brain is an electrochemical device. So, we were thinking of an electrochemical device. And Hoff came up with the idea of electroplating. So, if you have a plating solution, and you put a resistive element in the plating solution, you can plate metal down on top of this element and make it more conductive. So, you can change the conductivity of it by depositing metal or reversing the plating current and stripping it, taking metal from this substrate, this resistive substrate, and putting it back onto a source of metal. So, what we were using was copper. The first memistors, I went to the Stanford bookstore with an ohmmeter. And I put the ohmmeter down on the counter. And the young lady behind the counter said to me, "Can I help you, Sir." And I said, "Yeah, I'd like to buy some pencil leads with the highest electrical resistance." She said, "I beg your pardon." Then I showed her. And she brought out some pencil leads. And I put clip leads on the pencil leads. And the pencil lead that I selected was Fine Line type H. It's a hard lead. Fine Line was a brand of pencil and pencil lead. So, we took that, measured it. It measured nine ohms. And then we connected wires to the ends of it, insulated wires, and brought this outside the solution, a copper plating solution, which is copper sulfate and sulfuric acid. And then we put a copper rod inside the solution and were able to plate copper down onto the pencil lead. And it went from nine ohms down to about a quarter of an ohm. So, it had quite a range of resistance change. So, all the learning took place by changing the amount of copper deposited on the pencil leads. So, the first Adaline that we made with memistors was a pencil lead machine. Now, the machine is all gone. We don't have it. But what I've got is pictures of it. So, I can't show it to you today, but we do have pictures of it.

Hodge: Then what were your first implementations on a digital computer?

Widrow: These came later. What we did is eventually we got an IBM 1620 computer. This is a computer that was made in San Jose. And we got our first computer, maybe it was a year after Hoff and I first did the LMS algorithm. And we started doing all simulations of neural networks on the computer because what we physically built was a single neuron. Now, we wanted to work with a whole network of neurons. And it was very, very slow to do this on the computer. But we were able to do it. And that's how we got started with computer simulation of neural networks.

Hodge: So, the LMS general filtering technique got picked up fairly widely, probably by a lot of the people who weren't necessarily thinking of neural networks or learning machines. What was the idea—?

Widrow: How did that happen? Well, there's a man named Robert Lucky who is very famous in the annals of the IEEE.

Hodge: He writes a column, right?

Widrow: He writes a column [in IEEE Spectrum]. And Lucky developed the adaptive equalizer at Bell Labs. He had developed a learning algorithm, which is different from LMS, and, in my opinion, not as good. And when they—it's very hard for anything to ever get into Bell Labs because they're like MIT. The sun rises and sets right there. And all good things come from there. Anything coming from somewhere else is either suspect or ignorable. But somehow LMS got in there and got into their work on the adaptive modems. And it was the work on adaptive modems at Bell Labs that really caused it to get into digital communications.

Hodge: About what year was that?

Widrow: Oh gosh, I don't know fifty, sixty—probably '65, 66', '67. You'll have to Google and look up Lucky's first paper.

Hodge: Okay.

Widrow: That would be his first paper, but he shows adaptive algorithms different from LMS. When they actually started using LMS? That I'm not sure. But they did adopt LMS. And that's the way everything has been done ever since.

Hodge: Another paper you have mentioned in your CV that became something of a classic was one from December of '67 in the proceedings of the IEEE, "Adaptive Antenna Systems."

Widrow: Yeah, that became a citation classic.

Hodge: What was that paper about?

Widrow: Okay. That paper—I was on sabbatical. I started writing the paper before I went to Europe. Then I went to Europe. My students were here still working on it. So, that the manuscript kept going back and forth across the ocean by airmail. So it took two weeks [by mail] to go from Europe to California, two weeks to go from California back to Europe. But that's the way things were in those days. That paper was written and submitted to proceedings of the IEEE. And the editor had a dilemma. I can tell from the letter from the editor where he enclosed all the reviews of the paper. What the reviews were—one reviewer said reject, absolutely reject it. Another reviewer said publish it but make some changes. And he suggested some things. The third reviewer said I can't really make up my mind whether this paper should be accepted or rejected. He said there may be some interesting things in there, but I really don't understand it very well. So, the editor had to make a decision. And he could have decided one way or the other because it was sort of not clear how his reviewers are reviewing it. So, he decided he's going to publish it. The paper was so original, so unusual, nothing like it. It makes it very difficult for reviewers because what we say is off the wall. In other words, it's something out of the blue. You don't know where it comes from. What it had to do with is how to build an antenna, an antenna that can receive signals from a signal source so you can steer—it had a beam just like an antenna. You can aim, position an antenna. You can aim it. So, this had—this was steerable. You can aim it. But if interference was coming in from a direction other than the direction you're trying to look, called the look direction, you'd like to get rid of that interference. And you may not know which direction the interference is coming from. Now, an antenna that is directional is not perfectly directional. It has a main lobe, main direction of sensitivity, the look direction. But it also has sensitivity in the other directions. It has so-called side lobes all around it. Every antenna has that. I knew all about that because I was teaching that stuff in the class I taught on principles of radar. Actually, we developed this for sonar. We had sonar data from the Navy. And we figured that the Navy, if we called it adaptive hydrophone arrays, the Navy wouldn't let us publish it. So, we called it adaptive antennas because we knew our sponsor, the Navy, didn't care about antennas. They cared about hydrophones. So, we called it—instead of calling it adaptive hydrophone arrays, we called it adaptive antennas. So, we're talking about antennas. And so, who are reviewing it are antenna people, not sonar people. So, they're looking at this, and we're talking about changing coefficients. We're talking about an antenna has a whole series of elements, receiving elements. And each element is connected to a linear combiner. So, you're combining these signals and adding them together. Now, a signal coming from the look direction, if you just simply add the signals together, that reinforces the reception in the look direction. Other directions if you added the signals together, they wouldn't combine to totally reinforce. In some directions they can actually combine to cancel each other. So, an antenna has a main lobe. It has side lobes. And between the side lobes there are nulls where from those directions the response is zero. So, what we're trying to do is put nulls in the directions of noises, interference, not knowing where the interference is coming from. And we developed a way to do that with an adaptive algorithm that the antenna could learn to become a good antenna, to look in the look direction as directed, but also in directions where noises are coming to automatically put nulls. So, we were able to get tremendous performance from antennas that would be unheard of. And the ideas of what we were doing were so strange and so unusual to antenna people, if they understood it, they didn't like it. They didn't this. They didn't that. The paper became a classic. In other words it came that close to getting rejected. And that's the kind of paper that becomes a landmark paper.

Hodge: It sounds like what you proposed was fairly computationally intense.

Widrow: Yes.

Hodge: So, at the time if it was something mobile, it probably could actually be implemented.

Widrow: Ah, but then we published that paper. And we stopped working on that. We start working on other things. And meanwhile, what happened is that because we published it in the proceedings of the IEEE called adaptive antennas, people working on antennas picked it up and started working on it. And then some years after we published that paper, maybe ten years something like that, I received a phone call from somebody who was organizing a conference in Washington on the subject of adaptive antennas. I was flabbergasted. I couldn't believe it. So, they invited me to come and give the keynote speech at this conference. There were hundreds of people there. And they—people from companies—industry. And they showed their work. I just couldn't imagine that they were doing LMS algorithm at microwave frequencies. How the hell they could do it that fast? How they were able to do it, to develop circuits that can do it? But they did. You know when you're determined to do something you can do it.

Hodge: Was it analog circuitry, or was it—?

Widrow: Most of it analog. You couldn't make digital go that fast.

Hodge: Nowadays, is that basic technology present in most wireless communication devices?

Widrow: Wireless communication devices all have that. They—not adaptive antenna, but they have adaptive equalization and adaptive echo canceling. The next several more generations of cell phones are going—the cell phone is going to have a little antenna array. The cell tower will have antenna arrays. And they both will use adaptive antennas. If you can do that you can have in the same cell you can have somebody with a cell phone speaking in a given frequency channel, somebody else near him speaking in the same frequency channel at the same time. And what will happen is that your antenna in your cell phone will form a null toward the other guy using your same channel so that you can greatly increase the capacity of the cellular network without building more cell towers. So, guess what's going to happen? Now, you have to realize how long ago that paper was published.

Hodge: Right.

Widrow: But this is the technology—versions of it, modern versions with modern circuits to do this. That's what I say. Sometimes these things are so far ahead.

Hodge: In this case about fifty years between—

Widrow: Well, that's not untypical.

Hodge: How are—just out of curiosity, in either today if you're doing it, or the next generation of cell phones, how would you physically implement these? Would there be a dedicated processor for just that function? Would it be subsumed into something else?

Widrow: I would have no idea. But the chipmakers do magic.

Hodge: So, going to that idea of a chipmaker, I want to jump back a bit. You came out to Stanford right around the time that Shockley came out here to form Shockley Semiconductor. Did you know about that when you came out here? Did it factor into your decision?

Widrow: Well, I was—I switched from hardware, semi-conductor hardware. I switched from that to algorithms. So, that wouldn't have had any effect on myself. What had an effect was that John Linvill was here. And he was the one that brought me to Stanford. So, what—my first year here there was an IEEE conference in Denver, Colorado. No, it was in Boulder, Colorado. That's the first time I'd been in Colorado. John says I'm going to go to this workshop, an IEEE workshop on integrated circuits. And he was interested in the adaptive neurons because he thought that these things are going to become integrated circuits. And he was looking for good applications for integrated circuits. That's what got him very interested in the work that Hoff and I were doing. So, he said to me, "Would you like to come—I'm going to this workshop. Would you like to come with me? We'll go to Boulder and attend this conference." So, I went there. There were about a hundred people at that conference. Half of them thought the future of electronics is integrated circuits. The other half thought integrated circuits are not going to happen.

What was the reason? The reason was that in those days making transistors, the yield, the percent of the transistors that they made turned out to be good with something between five and fifteen percent. So, they say, “Look, we’re trying to make a transistor. Most of the transistors we make are rejects. We only get a very small yield. You’re talking about putting five transistors on the same substrate and having all the circuits interconnecting them made at the same time that the transistors are made. If a single transistor—if we only get ten percent yield, if you put down two transistors on the same chip, you’re only going to get one percent yield. You put five transistors on the same chip, you’re going to get nothing. It’s not going to happen.” Where they were wrong is this. If you put down a transistor on a substrate, you can put another one right along side of it. And if you get ten percent yield on one, whenever you get one that works, the one right next to it is also going to work. And that’s what made integrated circuits possible, that the yield was low, but if the process was right, and you can make a single transistor that works, there’s a very good chance you can make a whole bunch of them that work.

Hodge: This wasn’t known as Silicon Valley at the time when you first came here, did Shockley Labs and then Fairchild, was that something that was in the news, something that was discussed around campus, or when did this start?

Widrow: Oh, yeah. We all knew about Shockley Semiconductor. Yeah, they were located on Page Mill Road. And the big famous companies were Hewlett-Packard and Varian. And also Lockheed had a research lab in the Stanford Industrial Park. This was the beginning of the Stanford Industrial Park. This was long before Silicon Valley, so Silicon Valley developed as a result of all this, and half of it came from spinoffs from Hewlett-Packard.

Hodge: It’s often talked about now, the guys who left Shockley to form Fairchild was a very important turning point in the valley. Was that even something that was really noticed at the time contemporaneously?

Widrow: Oh, of course it was. Absolutely. See, the ringleader of the group was Bob Noyce. Bob Noyce was a good friend of my late colleague, Jim Angel. Jim Angel and I worked very closely together. Through Jim Angel I met Bob Noyce and Gordon Moore. When Intel was just getting started Ted Hoff already had his doctorate. Bob Noyce came to Jim Angel and said, “Are there- are there any students at Stanford who are getting PhDs that you could recommend, good scientists, good engineers, th- that- that- that would like- like to work with us?” And Jim Angel recommended Ted Hoff. But Ted Hoff was a postdoc in my lab. I just told Ted, I said, “Ted, this looks like a great opportunity.” I said, “You really have to do this.” I hated like hell to lose him. He was my right arm. I said, “Go and do this,” and he did. And when he went to Intel he became employee number 12. And not long after that, well, that’s a whole separate story, but Japanese people came with design of a desk calculator that they were going to build. They wanted a chip set. They wanted Intel to make a chip set, so Bob Noyce sent the Japanese guys to go see Ted Hoff. Ted told Bob Noyce, “They have a very good design. It’s gonna work. It’ll be fine.” But he said, “Look, if we do this for them we got other people coming in with another design and another design.” He said, “Why- why don't we make a chip that is programmable and we can program it for all of our customers so we don't have to design a custom chip for everybody that comes?" So Bob Noyce liked the idea. Gordon Moore liked the idea. They gave him two people to work with, so he had a group of three people. He sat down and designed the whole thing. And then, let's see, Federico Faggin designed the software, the operating system. And another man, a third man, I forgot what his name is [Stanley Mazor], that did the chip layout, and they made the first microprocessor.

Hodge: So how did a PhD student who was a signal processing specialist wind up becoming a microprocessor architect?

Widrow: I'll tell you how. When we had the computer in the lab, that IBM computer, if there ever was a problem with the computer we'd call the guys in San Jose and they'd send a guy to fix it. They called him the CE, the customer engineer. When the CE came to fix the computer all the circuit boards were on trays. They slid out. When the guy came in, my students would pull out the tray where the fault was and tell him practically which card to replace. They knew so much about the insides of that machine. The IBM 1620 was what's called a BCD, binary-coded decimal machine. It did all the arithmetic in decimal, but it used four bits to do each decimal digit, so memory and logic, everything, was binary-coded decimal. Now, of course, everything is done in binary. You start out in decimal, convert it to binary, do everything in binary and then convert it back to decimal. But then they did everything in decimal. You know, these were raging fights. How should you build a machine? Should it be all binary or should it be BCD? That machine was BCD, so the basic memory, chunk of memory, was four bits. The first Intel microprocessor was a four-bit machine, so a lot of the stuff that Ted had learned playing with that computer. Now, among all my students over the years, they all worked on neural nets, adaptive filters, all kinds of problems in signal processing, et cetera, et cetera. Half of them went into the computer field. The other half stayed working on adaptive this or that or signal processing this or that. So only half of them really pursued careers where they were using directly stuff that they had done in their doctoral research. The other half got involved with computing somehow or other, and we had computer in the lab, and that explains it. But there's another aspect. I'll tell you a little bit about Ted Hoff. He's one of the smartest guys I've ever known. Another guy who was one of the smartest guys I've ever known was William Shockley. These are people with such a breadth of knowledge and such an amazing imagination. They're truly unusual people. Another one I met who was an incredibly smart guy was a physicist at Caltech, the late Richard Feynman.

Hodge: Pardon me?

Widrow: Feynman, Richard Feynman. Brilliant guy. What an amazing man. So these are some people I met that I think are just truly exceptional. I put Ted Hoff up into the race with the Nobel Prize winners.

Hodge: I'm sure he'd be happy to hear that.

Widrow: Don’t tell him I said it.

<laughter>

Hodge: Now, Shockley, you knew him? You personally knew him then?

Widrow: He was a colleague. He became faculty member in Electrical Engineering. When Shockley Semiconductor folded he joined the double EE faculty at Stanford. So he retired from Bell Labs and then Shockley Semiconductor. Then he retired from Shockley Semiconductor, became a Stanford faculty member. Eventually he retired from Stanford faculty, but he lived on the campus. I lived on the campus. My house and his house were adjacent, so we’d be out there in the fall every year raking leaves together. And when I had a technical problem about my memistor, about the chemistry, making it better, he was the guy we went to talk to. He had some—

Hodge: Was he helpful?

Widrow: He had some good suggestions. He caught onto the idea like zoom! And he knew all about chemistry. Ted Hoff was the same way. Knows about physics. He knows about chemistry. There’s just nothing scientific that the guy doesn’t know about. Ted is like that. Shockley was like that. Shockley knew everything, and Shockley made the big end of career mistake of getting involved in social problems. And a social scientist said he doesn’t know what he’s talking about, but the social scientist didn’t know Shockley. So I think I’d better not make any comments from hereon because the stuff was so controversial, and I’m not sure whether he was right or wrong.

Hodge: Yeah, he certainly had his late—

Widrow: But I would certainly not dismiss him out of sight on anything, anything that he studied. Excuse me.

Hodge: No, that’s fine. I was going to say that his late career certainly had a lot of controversy to it. Prior to that stage when he was an EE professor at Stanford was there anything particularly interesting in his work at that time after Shockley Labs?

Widrow: I don’t really know whether he came up with anything really remarkable as a faculty member. But he was teaching classes in semiconductor theory, stuff really for physicists and engineers. He was sort of halfway between.

Hodge: Okay. Then another question branching off. Since you had Ted go to Intel and start working on microprocessors at any point did you start considering working with someone to get a special purpose processor to implement your algorithms?

Widrow: We've always thought of that but the idea of making a custom chip to do it is very expensive. It's not something you could do. You had to have a lot of money behind the idea. And anybody that's going to put a lot of money into it has to have some idea of what they're going to get out of it. They don't do it for the science of it or the fun of it. There's no doubt you could do it. I mean, we're simulating neural networks on the computer, so it seems to me that there's no doubt you can build special chips, and we're even contemplating this today.

Hodge: Yeah, so starting in the late seventies you started to see the appearance of dedicated digital signal processors on the market. In fact, Dr. Hoff, I believe, designed one while he was at Intel.

Widrow: He designed the first DSP chip, digital signal processing chip.

Hodge: Did you start using one in your research out of curiosity?

Widrow: It would've been interesting to have it and I knew he did it, but we didn't have real-time applications that we needed a chip like that for. Anything you could do with that chip you can do with a conventional computer. So that's the way we were doing things without special purpose stuff.

Hodge: Looking back on the work you've done in neural networks, adaptive learning, have they fulfilled the promise that you thought they had?

Widrow: In some sense yes. In some sense no. Some of the commercial speech recognition systems now are using neural networks, so I think there are applications in neural networks that are performing tremendous human good. Also, the other thing is that the adaptive filter is like a neural net with a single neuron, and without adaptive filters you wouldn't have digital communications. So do I say is this disappointing or not? I say the whole thing didn't work out the way I thought it would, but I ain't complaining.

Hodge: <laughs> How much interaction was there between you and computer scientists who studied AI?

Widrow: Zero.

Hodge: Okay. So do you have an opinion on sort of strong AI versus weak AI, how smart computers will get in the future, or would you like to not predict that?

Widrow: It's interesting subject, interesting subject. Some of the best work that was ever done in AI was done by a man named Arthur Samuel, and what he did is he developed a program. He was working at IBM at the time, worked the program using an IBM mainframe, to play the game of checkers. And he was able to get his program to play a game of checkers that he could not beat the program. Program would beat him, but it wouldn't beat the best checker players in the world. But the stuff he did corresponded to some thoughts I had about thinking, how thinking works, and this is things I thought about after I came back from the conference at Dartmouth College. Samuel's work was so long ago. Samuel's work was before the conference at Dartmouth College, and he was already working on checker-playing program. I think he published a nice paper in 1959 on checker playing. It was in "MIT Technology Review" magazine. I still have a copy of it. It's a beautiful article only a few pages long, but really, really profound. And to me that's the best stuff ever done on AI. Other things done over the years in AI have turned out to be very useful. They don't really, I think, reflect anything whatsoever to do with human thinking. Stuff I'm working on right now I didn't tell you about yet, but what I'm working on now is human memory. How does it work, and how do you build a human-like memory for a computer? And I think I know how to do it. We've already developed pieces of it that are working fine. And I think that this is—I would put that work in the field of AI. See, there's an AI community but then there's a much bigger community working on various aspects of AI that don't put themselves into the AI mafia. The AI mafia is centered around DARPA support. DARPA has spent hundred of millions on AI, and I hardly ever got a little snitch of it. So they don't like the way I do things. So the way I do things is different from the official AI community. But if you talk about an adaptive filter that learns this is long, long, long, years before, decades before so-called all of the sudden the AI community discovered learning, machine learning. Now they talk about machine learning. The first guy to use the words machine learning, to my knowledge, was Arthur Samuel way, way, way back, long before ADALINE. So there's an official AI community, and this exists at Stanford. You know, people teaching machine learning. They have discovered things that people in my gang, not me necessarily but other people that I have worked with, have contacted with, have developed like learning and other things like that back in the 1960s. These guys are rediscovering it. They're doing it better. They've got much bigger computers to do it on than we had back in the sixties. I can't say that there's nothing new and original in the work they're doing on machine learning. They probably have done a few bits and pieces here and there that are original, new and productive. But it's unfortunate that there's no real connection between these two communities. People working on learning for years, long before the AI community discovered learning. They were doing all kinds of artificial intelligence but no learning. How the hell do you build something that's supposed to be artificially intelligent without learning? Then finally they discovered learning. They gave it a buzzword, "machine learning." We never had that buzzword. We just called it learning systems. So, you know, you ask a sensitive question and I can't explain it. But I look on their work. I see what they're doing. I say, "Yeah, this is good. This is fine. It's okay." They look on what we do as <makes ptthhh sound>. The only thing is we were there 50 years before they were.

Hodge: One thing you mentioned earlier in this interview was you went on to adaptive systems after doing work in quantization noise. You said you could've spent the rest of your career doing quantization noise. But you did actually come back, and it seems like out of nowhere wrote a book on quantization noise fairly recently.

Widrow: Who did?

Hodge: You did.

Widrow: We did. That's right.

Hodge: What caused you to go back to it after so much time?

Widrow: Well, I'll tell you. I hadn't intended. I felt when I was satisfied with our work on neural nets I was going to go back and ultimately do a book on quantization noise. But I started on that book long before I thought I would, but long, long after I had stopped working on it. What happened was my coauthor is a man named Istvan Kollar. And Istvan is a professor in a university in Budapest. He has a brother who was working with a colleague in the aero and astro department at Stanford, and Istvan's brother was on sabbatical. The two of them are professors at this university. And Istvan was visiting his brother. His brother was at Stanford. Istvan, I think, was in the country to give a paper at a conference, came to Stanford, was wandering around in our building and he saw my name on the door. And he got all excited because he had a group of students in Budapest working on quantization noise, and the basic reference, the beginning of it, was the work I wrote right after I did my doctoral thesis in 1956. So they were working on this stuff. I didn't know anything about it. And I met him and the two of us bonded right away, and we decided we're going to do a book. And he said, "I need to come back to Stanford." I says, "You need to come back to Stanford, that's right." He said, "Write- write a letter and I can get a Fulbright. So I wrote a letter and he got a Fulbright, so he came for a year. I pushed aside everything, just worked with Istvan. Of course, I had students, doctoral students, in the pipeline and so I helped them when necessary, but I stopped working on things that they were working on and focused everything on working with Istvan. So day and night I'm working with Istvan. So after a year we're not done yet, so we wrote another letter and he got another extension on his Fulbright for another year. Meanwhile, his university is getting ants in the pants because they want him back there to teach. So after the second year he had to go back. We're still not finished. So it took about ten years of back and forth in mail, in email and phone calls to do work that we could've finished it up in a couple of months. But we did. We finished it, got it done, and it now exists as a volume on quantization noise. It's the only one of its kind.

Hodge: Well, if you're going to continue in this theme you're now going to have to do a book on core memory.

Widrow: <laughs> No, you don't do a book on it. Core memory is history.

Hodge: <laughs>

<crew talk>

Hodge: So one of the other things I wanted to ask you about is you've been teaching engineering for quite a while now. Have there been any major changes? One of the changes you've talked about is that there were only 11 female students when you entered MIT. That's obviously changed. But has there been other changes in the way engineering's taught over the last 50 years?

Widrow: Well, the subject's quite different. You know, electrical engineering, the field evolves at a very rapid pace, and the way the teaching works at graduate level, students taking graduate courses, the graduate courses are taught by faculty who are doing research. So whatever you discover in the research lab goes immediately into your coursework, into your curriculum. So it's a constant changing, adding new things, and of course, things that turn out not to be as interesting you sort of shy away and bring in new stuff. So this is a constant change. Undergraduate classes change much more slowly, but graduate courses change with the research.

Hodge: Is the level of specialization that students come out with narrower now than it was when you graduated?

Widrow: Doesn't seem to be. We have a bigger faculty. You know, when I first came to Stanford the electrical engineering department had something like 15 faculty. When I left MIT at that time there were 60 faculty in electrical engineering, so I went from a department with 60 faculty down to a department with 15 faculty. The difference was going from MIT to Stanford was that every single piece of electrical engineering you had several faculty member working in that area. Here at Stanford in each part of electrical engineering you had one person. Everybody counts. So even being an assistant professor, a little baby assistant professor, the first day he walks in the door he's got doctoral students. If I were at MIT I wouldn't have a doctoral student for years. Doctoral students would only go to the most senior professors. So there was a sort of equality between the youngest assistant professor and the most senior full professor at Stanford that didn't exist at MIT.

Hodge: We've covered a lot of ground in this interview. Was there anything that you wanted to get in that we might've skipped over or skipped past before you got to say it?

Widrow: Well, we've covered all sorts of things.

Hodge: So we're running out of time here for today, but there is as you brought up to me during the break we haven't talked a lot about sort of the contour of the neural networking field. As you mentioned, you probably were there at the very start, at the birth of the field. So do you just want to give sort of your thoughts, your overview of where you think it's going, what the biggest accomplishments were? I realize that's a very open-ended question.

Widrow: Most of the applications, aside from what goes on in adaptive communications, in the field of neural networks most of the applications one way or the other relate to pattern recognition. And about 10 years ago I realized I'm coming close to the end of my career and I worked in pattern recognition and adaptive control systems and all sorts of learning systems for years, but the pattern recognition hasn't really progressed. It hasn't really advanced as I thought it would, and I was disappointed. And I was saying to myself, "How- how does the human do it? How does the animal do it? Why do we do it so well compared to whatever we're able to be- whatever is able to be done by a computer or by artificial means? What are- what are we doing that's wrong?" So I decided to take everything I've done on neural nets. That's really not the way to go for pattern recognition. And I just made a strong statement. So I said, "How- how do people do it? What do people do? How does that work? How are you able to recognize things? How are you able to recognize sounds, sight?" And I began to think that this has to do with memory. Memory is the key. When you try to recognize something that's in your hand you're looking at it. You're trying to recognize it. The way you do that, it seems to me, is to compare what you're seeing with something you've seen before that you have already stored the image of in memory. So the way you do pattern recognition with neural networks directly is you feed in input patterns and you train the neural net to give outputs corresponding to the class of that input pattern. You teach it to classify the input pattern. Then you give it a new pattern that'll tell you the class of that pattern that's closest to one of the patterns that you've already seen. But the patterns that you've trained it on, once you've trained the neural network you can throw the patterns away. All the knowledge, all the experiences are contained in the synapses and the weights, in the coefficients. Now, what I'm saying is don't throw away the training patterns. Keep the training patterns, and what you need to do is to connect something you're seeing or hearing or feeling or smelling or something now to something you've experienced previously in the past. And you have that all in memory. What I think is everything that you encounter, that you experience, that you see, hear, smell, feel, whatever that's interesting to you is going to be stored in the memory in your head for the rest of your life. It's going to be there forever, as long as you live. The problem is how do you retrieve things from memory when you need to? Let's say you're trying to recognize a person's face. Let's say you see a crowd of people, hundreds of people. All of the sudden in the crowd you see a face that's a familiar face and you say, "I know that guy." How do you do that? "I know that guy." And you go up and talk to that person. You can have a conversation with that person because you know the person. Clearly there has to be things that a previous conversation or conversations about that person that you remember. It's in memory. If you didn't have memory you wouldn't know that person. You wouldn't be able to have a conversation with that person. So with seeing that person you're able to connect the image of that person's face even though you never saw it exactly the same way, same distance. If he's further away the image of the face on the retinas of your eyeballs is smaller. He's up close, bigger. So you see his face different sizes. You see different aspect angles. The head is this way. The head is that way. Your head is this way, that way. So the way you see the person's face is different every time you see the person. What I say is I'll have a conversation with you for a half hour. I get introduced to you. We introduce each other, however. I see images of your face, subject of conversation, your name because we were introduced. The next time I see you, if I can pull that out of memory the images of your face, the sound of your name being pronounced, they're all recorded in the same place in the memory. You don't have a special memory for faces or a special memory for visual images distinct from sounds. They're all recorded in the same place. They have to be. I'm trying to think. If I were going to design a memory, and I did design memories. I used to design core memories. This is different, but I'm still a memory designer. If I'm going to design a memory to behave like a human-like memory, how would I design it and what would I do with that memory if I had it? Now what I realize you can do with it if you had a computer with a human-like memory you'd be able to recognize patterns like humans. You'd be able to do very complex control problems like the human does with locomotion. That's all done from memory. Without memory you can't walk. Without memory you can't talk. The control is all pulled out of memory, so the connection between what you see now and what you heard before or what you've seen before, what you've seen before is in memory. You have to connect these. And I use neural nets to make the connection, to do the association. That's what I'm doing now. That's my current research.

Hodge: Well, it's good to hear that you're still actively researching this.

Widrow: I retired several years ago. I still have a half-dozen doctoral students and I just finished teaching a class. I teach one class per quarter at Stanford at this point, so I know it's a light load for me.

Hodge: <laughs> Well, it's fitting that we began this discussion with memory and ended it with memory, and on that note as well our camera is out of memory.

Widrow: <laughs>

Hodge: So Dr. Widrow, thank you for your time and thank you for coming here to talk to us today.

Widrow: It's a pleasure. Thank you.

END OF INTERVIEW

