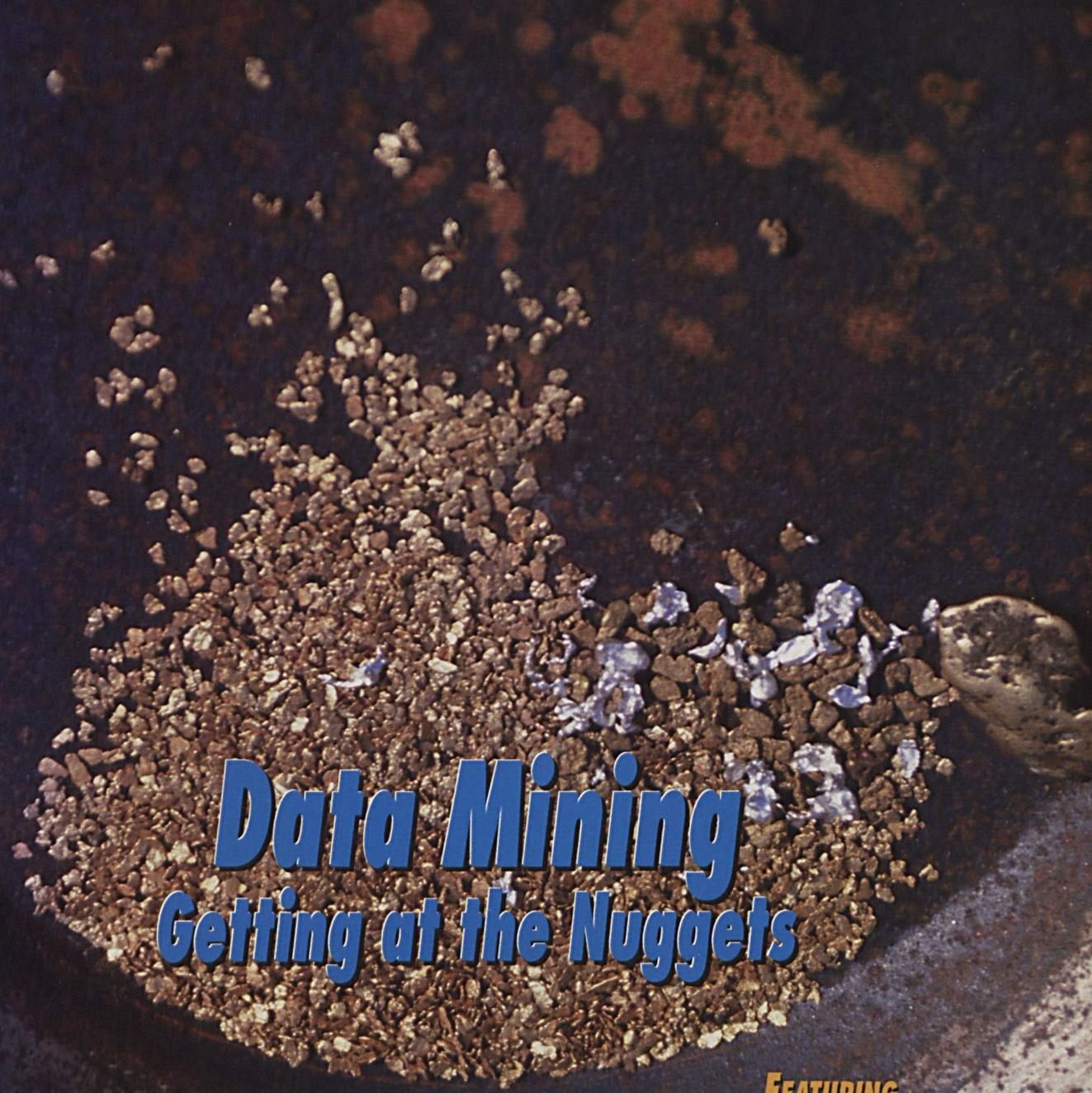


OCTOBER 1996

IEEE EXPERT

# INTELLIGENT SYSTEMS & THEIR APPLICATIONS



## Data Mining Getting at the Nuggets

### FEATURING

*A conversation with AI pioneer  
Oliver Selfridge*

*Using a knowledge-based architecture  
for a design automation application*

*Facing Web content  
distribution challenges*

*What makes a compelling  
empirical evaluation?*

 IEEE  
COMPUTER SOCIETY  
50 YEARS OF SERVICE • 1946-1996

 THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

# Oliver Selfridge— in from the start

Peter Selfridge  
pgs@research.att.com

Driven by his curiosity about the nature of learning, Oliver Selfridge has spent over a half century enmeshed in the most exciting developments in artificial intelligence, communications, and computer science. A participant at the original conference at Dartmouth in 1956 (and at the Western Joint Computer Conference in Los Angeles the year before, which he considers the true start of AI), Selfridge formed working relationships and cemented friendships with AI's founding members—John McCarthy, Marvin Minsky, and Allen Newell, among others—as he went on to become a true AI pioneer himself.

Before retiring in 1993 after 10 years as Chief Scientist at GTE Laboratories, Computer and Information Systems Lab, he served as a member of the National Security Agency's Advisory Board for 20 years, chairing its Data Processing Panel for the last 15 of those. He also served on various advisory panels to the White House, as well as on the peer review committee for the National Institute of Health (NIH), directed Project MAC and the Cambridge Project at MIT's Lincoln Labs, and was Staff Scientist at Bolt, Beranek, and Neuman (BBN). He continues to write and speak on machine learning and AI, and especially on self-improving systems.

This profound, long-term familiarity with both the philosophical underpinnings of AI and the practical application of AI technology places him in a most advantageous position to comment on the growth of the field in this 40th anniversary year of the Dartmouth Conference. Peter Selfridge, Oliver's son and an AI luminary in his own right (AT&T Laboratories—Research, previously part of AT&T Bell Laboratories, member of *IEEE Expert's* Editorial Board), recently asked his father to assess AI's progress since its earliest days.

*Peter Selfridge: How did you become interested in AI?*

*Oliver Selfridge:* It was at MIT, a long time before the Dartmouth Conference, and

I was studying mathematics under Norbert Wiener. By luck, of which I've had a great deal in my life, I was introduced to Walter Pitts, who was working with Warren McCulloch on a topic they called *theoretical neurophysiology*. I had studied logic, and through Walter, Warren, and Norbert got introduced to neural nets at that time. I went to the Pacific at the end of World War II with the US Navy and came back to graduate school, again at MIT. Norbert was then writing *Cybernetics*,<sup>1</sup> and Walter and I were helping him with various aspects of it. As I studied mathematics (my original field) and interacted with Norbert, Warren, and Walter, I began to be interested in the specific processing that neural nets could do and even more interested in the general properties of learning.

At this point McCulloch and Pitts had written the first two AI papers (although it wasn't called that).<sup>2</sup> The first showed that a neural net could work out certain kinds of problems, such as pattern recognition in the general cognitive sense, and the second discussed acquisition of patterns (how we know "universals"). These two works followed all the glorious mathematics that Turing and Gödel had done in the twenties and thirties about computability and Turing machines. This mathematics was, of course, the beginning of a formal description of what computability meant. Johnny Von Neumann visited us at MIT occasionally, so again by pure luck, before the age of twenty, I had been introduced to McCulloch, Pitts, Wiener, and Von Neumann.

*Your original background is mathematics, but many of these ideas begin to verge into biology. Was this accidental?*

I went through MIT, and so I got a very fair liberal arts training, including things such as biology. One of my roommates was Jerry Lettvin, a super neurophysiologist, and McCulloch had also been a neurophysiologist of some renown.<sup>3</sup> When one looks at the early work on neural nets, you find yourself interested not just in the computing

powers of simple processing units but also in the physiology of computing in the brain.

*What was the next stage for you?*

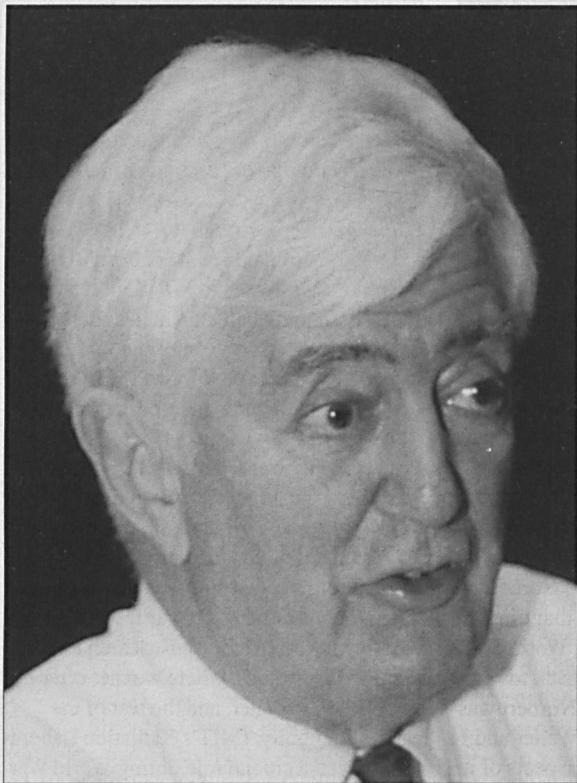
In 1951, I transferred from Fort Monmouth in New Jersey to MIT Lincoln Labs (Fort Monmouth had knuckled under to Senator Joe McCarthy, but Lincoln Labs had not, and therefore I moved). By another stroke of luck, I bumped into Marvin Minsky, who was fresh from a Doctorate at Princeton and was a Fellow at Harvard. He was a very bright mathematician and thinker in general—his thesis was on a particular model of, again, neural nets. Marvin worked for me one summer and then got a job in the mathematics department at MIT, working next to RLE, the Research Laboratory for Electronics, where we met often with Warren, Walter, and the rest of the early community. (MIT's Radiation Laboratory played a crucial role during World War II. It invented and built all of the really powerful radar systems in the latter part of the war. It evolved into RLE, which is celebrating its 50th anniversary this fall.)

At that point, we were using the term *artificial intelligence* and thinking about many of the general ideas about intelligence that are still around today—the tough cognitive processing that people did, perception, games like chess. We were trying to tie in the modeling that could be done with neural nets to these more obvious intelligent behaviors of people. My stress was then, as it still is, to understand how learning takes place as the primary source of intelligence in people.

*This was before computers were generally available—true? Was it before Von Neumann had invented the programmable computer? If so, how did you work—on paper?*

We were essentially doing mathematical and algorithmic modeling on paper. This was in 1953, so we knew all about Von Neumann's work—he had visited us and understood what we were up to. And, while we weren't writing AI programs yet, we had had some experience with early computers. I had done some programming on Worldwind, a very old computer put together by Jerry Wiesner at MIT to get money out of the Air Force (which it did superbly well). Lincoln Labs was a follow-on to that.

So, while Marvin and I were having these interesting thoughts, computer technology was exploding in the basement of



Oliver Selfridge, 1996.

Building C at Lincoln Labs where they had the Memory Test computer, all vacuum tubes, and then the Transistor Test computer (TX0). Jay Forrester was at MIT, and he came out and perfected the donut-shaped core memories, and I remember seeing the first million-bit memory. Now a million-bit memory is 64K words or 128K bytes, which, if you had it on your fingernail now, you would hardly be able to see it. This million-bit memory occupied, with its frame, about one cubic yard. The power supply, built with vacuum tubes, took another cubic yard. Marvin and I speculated about whether commercial memory might ever be as cheap as a dollar a bit! This was only 40 years ago.

The programming was usually done on paper tape in machine code, essentially. And at about this time, IBM came out with the 650. The 650 didn't read a tape, but instead, you programmed the machine with wires and plugs, much like the old switchboards.

*Back to the early AI community. Who was involved at that time?*

There were numbers of us thinking quite hard about AI. John McCarthy had met Marvin at MIT. John was the one who got the funding for the Dartmouth Conference from John Morrison who worked at the Rockefeller Foundation. There were numbers of other people also beginning to be

interested—Wes Clark and Belmont Farley at Lincoln Labs come to mind. Belmont Farley wanted to be a wet neurophysiologist, and so he was interested in neural nets, too. By luck, at one point I made a trip out to the Rand Corporation—this was 1954—in Santa Monica, California. Rand was smaller then; I was talking to Willis Ware and a young, smart psychologist from Carnegie—it wasn't Carnegie Mellon then, it was Carnegie Institute of Technology—named Allen Newell. I talked about pattern recognition. And Allen and I hit it off, and this was really the trigger. I mean he was ready for it. He got the whole idea, he got turned on, and he went back to

Carnegie and turned Herb Simon on. Those two went on to become, well, Newell and Simon of CMU, extraordinarily influential thinkers over the years.

Marvin and I were talking and writing about these things in general. What did we think the problems were?—visual pattern recognition, making speech, understanding speech. Later, *planning* played a bigger role, but not then very much. I still had learning very much in mind then, as I still have. Marvin and I essentially organized what to me is more the beginning of AI than the Dartmouth Conference, which was the Western Joint Computer Conference, held in Los Angeles in 1955, the year before the Dartmouth conference. We got funded by, I think, the ACM and the Institute of Radio Engineers, which I believe turned into the IEEE.

*Was everybody there?*

There weren't many "everybods," but Marvin, I, Belmont Farley and Wes Clark, and Allen Newell gave the four papers. It was a great thrill, realizing that this was a first, that this was going to be something—which it is.

*You were all in your late twenties?*

Marvin is a couple years younger than I (I was sure that he would catch up, but he never has). Allen was about Marvin's age,

and so yes, we were in our late twenties. There was a great deal of excitement.

*Can you now talk about the Dartmouth Conference, which people think of as being the start of AI?*

The Dartmouth Conference was funded by the Rockefeller Foundation, as I mentioned. It started, I think, on August 6, 1956, and lasted roughly four weeks. In that sense, it wasn't a conference like today—it went on much longer, and was much more loosely structured. Many people were invited and dropped by. I was not there the whole time. We would meet and give talks and argue, all those wonderful things. Many came who shared our goals and excitement: Roland Silver, Art Samuel, Leon Harmon. There were not many people doing research in computing, but there was lots of interest among those who were. John Backus, who invented Fortran, turned up there.

And it was a John McCarthy show, so to speak. At this stage, he was already beginning to think about symbol processing. Alan Newell really pushed us all into thinking symbol processing rather than bit processing—well, not quite *rather than*, but *as well as* bit processing. AI needed bit processing for, say, sensory integration for example, but if you wanted to do higher levels, you wanted to look at what it took to process symbols.

*Was John McCarthy on the West Coast at this point?*

No, John was at MIT with Marvin. I don't know John's years at MIT, but they must be something like 1954 to '58 or '59. After the conference, John was striding around his and Marvin's office inventing LISP. "CAR, CDR, CDR, CAR" echoing around. LISP, of course, became the preferred AI programming language for researchers, probably to this day.

*So was there the same sense you had the year before, that this was the beginning of something very exciting?*

It was tremendously exciting, both from the point of view of computation and from the point of view of beginning to understand what a mind is.

How does a mind do its computation? This is a tremendously antique question. By antique, I mean that it goes back 5,000 years at least, when people were looking at the clockwork of the heavens. Clockwork

is literally a copy of the rotation of the heavens, which is why up to a few decades ago clocks all had hands that went around—a model of the Big Dipper going around the pole star. You see lots of discussions early on about how to make a robot or a man model out of clockwork.

But here we were, beginning to get at these issues in computer software, where we had a chance of being able to find out the basics, to see what was do-able, and to get at the secrets about the nature of man.

*Did you all have the feeling that you were part of a grand tradition, spanning millennia?*

Marvin, John, Allen, and I, and many of the others, had considerable backgrounds, and we had studied philosophy. Not only that, but we also knew McCulloch, who was a philosopher of some depth. Warren was far more on the emotional side of the folks doing AI, and Norbert of *Cybernetics* was an emotional baby who had the math down cold. Wonderful, contrasting intellects to work with. Also at MIT then was Giorgio di Santillana, a Historian of Science, an exciting and visionary philosopher, a special hero of mine. It was very strange, as I said, that I had pure luck all my life with these things. But the conviction that this was very important, I still hold today—it is! We are facing the same questions that have been around for 5,000 years, questions about the nature of mankind. *What is man, that thou art mindful of him?* (Psalm 8, verse 4)

*Back to the Dartmouth Conference: what visions were discussed? Robots?*

We were really thinking about robots, talking about robots. Most of the problems were much harder than we thought. I remember a year or two after the Dartmouth Conference telling quite a good chess player (I believe it was Hugh Alexander, one-time chess champion of England) that my machine would beat him in five years. Well, it took 35 years, but now the world chess championship is clearly within reach. Because we understand them so well, people don't think of today's chess programs as AI. However, back then, the kind of look-ahead procedures and evaluations used were definitely part of AI. Ordinary computation didn't use those notions about planning, look-ahead, and tree searches. There wasn't anything like it in computa-

tion (ideas such as planning ahead for CPU processing or disk access); those were derived directly from interaction with AI and the AI techniques.

*Was the Turing Test discussed?*

Oh, it was indeed—thought of and discussed, as was the Turing machine. This was, after all, the breakthrough model of computation. It became then, and still is probably, a very powerful educational device—then and now, in advanced computation classes, homework assignments would often be to program a Turing machine. Back then it was recognized that there were interesting questions about which computational models were equivalent to Turing machines. At the time of the Dartmouth conference, there were certain mathematical games called Post tag systems.

***Because we understand them so well, people don't think of today's chess programs as AI. Back then, the kind of look-ahead procedures and evaluations we used were definitely part of AI.***

Marvin showed then that these systems could be interpreted as Turing machines. That is, a tag system can compute anything a computer can. That's not practically useful, but it gives a deep sort of foundation to understanding the basis of computation.

*Was this part of the beginning of complexity theory?*

Yes, but it wasn't thought of that way. Complexity theory itself took another 12 or 15 years before it had its name as such.

*What was the next step, now that computers were beginning to be available?*

There were lots of things discussed, yet we were handicapped by the inadequacies of computers. Some of the early stuff on neural nets worked appallingly slowly. Marvin and I recognized that computational capabilities had to be improved. Seven years after the Dartmouth Conference, again owing to a piece of wonderful luck, I was part of Project MAC, started at MIT.

This was the first true hands-on, widespread hands-on equipment. We were also lucky because a bunch of people from Lincoln Labs, where I was working (many of them my friends) went off and started Digital Equipment Corporation. DEC was building computers that were essentially run by users in those days. Before that, in the fifties, you ran a program by writing punch cards, which you then submitted to the desk of the computation center. The cards would be scanned and often sent them back with an error message such as "it doesn't parse" or it's no good for some other reason. If you were lucky, the deck ran after only a few iterations and you got your results back, but perhaps days or weeks later.

Well, when Digital built the PDP-1, you sat at the console and you wrote your program yourself and *you* could try it. This was a startling revelation. In fact, the first time-sharing system was done on the PDP-1. McCarthy at that point was spending the summer at Bolt, Beranek, and Neuman (BBN), with its Number 1 and Number 2 PDP-1s. BBN was an acoustical research center, but Licklider was there, J.C.R. Licklider out of the Harvard Acoustic Lab. Marvin and I both knew him very well. I had met him in the forties. He had been working on acoustics, but he was also wonderfully interested in computation, as a way to go by itself and for AI. He went down to Washington in the early sixties to run the Information Processing Technology Office, one of the major offices of ARPA, the Advanced Research Project Agency, for the Air Force. And Licklider gave some money to his old firm BBN to do something interesting, like time sharing. He played an absolutely critical role in funding AI work in those days, and he has been largely forgotten.

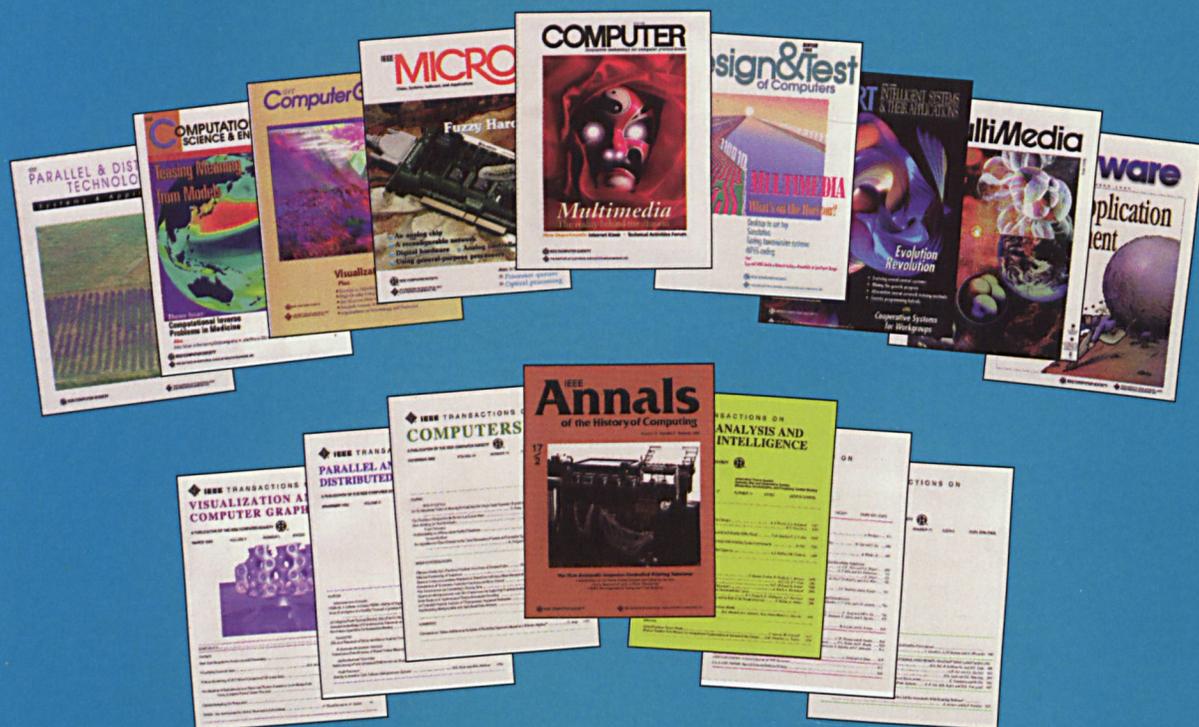
*When was time sharing first tried?*

McCarthy hacked time sharing at BBN in 1961 or 1962 on a PDP-1 with a total of 16K of memory, using not the old model 33 Teletype, which came later, but the Teletype that used a paper tape punch—click, clack, clack, clack. John had three users simultaneously—they weren't able to do much, but to do *something*.

This work continued at Project MAC at MIT, which was set up again by Licklider, who gave ARPA money to MIT in the early days. I should also pay credit to Professor

*Interview continues on p. 84*

# Chances are you missed some of these.



## Now's your chance to catch up.



Read the articles you missed. Find published descriptions of new products. Search for references or articles relevant to your research or current projects.

**Computer Society '95** is a fully searchable CD-ROM containing all 16 magazines and transactions published by your society in 1995.

**Special Member Price: \$89.95.** *Catalog #SW00049.*



Order from the  
IEEE Computer Society Online Catalog  
on our secure web site:  
**[www.computer.org/cspress](http://www.computer.org/cspress)**  
or call 1-800-CS-BOOKS