Avi Silberschatz Speaks Out

on Academia versus Industrial Labs, Startup Fever, Database PhD Quality, the Future of Data Mining as a Research Area, IP Issues, How He Wouldn't Change a Thing, and More

by Marianne Winslett

Avi Silberschatz
http://www.bell-labs.com/user/avi

This issue's interview with Avi Silberschatz took place in November 2001, in Roy Campbell's HDTV studio at the University of Illinois at Urbana-Champaign. Once again I'd like to thank everyone who suggested questions for this interview and for the other interviews in the series. As usual, all errors of transcription are my own and the videotape itself is the final arbiter of meaning. This interview was the first to be recorded digitally, another step toward our goal of making all the interview videos available on the SIGMOD Web site.

In upcoming columns, we'll be hearing from David DeWitt and from Hector Garcia-Molina.

==========

Welcome to ACM SIGMOD Record's series of interviews with distinguished members of the database community. I'm Marianne Winslett, and today I have here with me Avi Silberschatz, who is the vice president of the Information Sciences Research Center at Bell Labs. Before he joined Bell Labs, he had an endowed chair at the Department of Computer Science at the University of Texas at Austin. He's interested in operating systems, database systems, and distributed systems. He has also served as a member of the Biodiversity and Ecosystems Panel on President Clinton's Committee of Advisors on Science and Technology. He's been an advisor to the National Science Foundation and to a number of private companies. Avi is an ACM Fellow and an IEEE Fellow; he received the ACM Karl Karlstrom Outstanding Educator Award in 1998; he got the ACM SIGMOD Contribution Award in 1997; and he has an outstanding paper award from the IEEE Computer Society for an article in Transactions on Software Engineering. *And he's also one of the authors of two popular textbooks, one on operating systems and one on databases. So, Avi, welcome!*

Thank you.

Avi, you were a professor at UT Austin for many years, and then you took a position at a research lab, at Lucent. When you made the switch, did reality match your expectations?

Actually, I went there initially on a leave of absence for a year, and was hoping to go back to the University of Texas after one year. The main reason I wanted to go [to Lucent] was to actually build a large system, which you cannot do in the university.

> [T]he good thing is that there is no more [startup] fever because there are no more startups …

Ah.

After one year, I was in the midst of building this large system, so I could not go back to UT. In another year, I was still not finished with the large system, and after four years I decided that I liked it at Lucent and Bell Labs, and decided not to go back to the university.

So, then, your expectation was that [Lucent] would be a good place to build a large system.

It was an expectation which was met fully. I [was involved with building] several systems[,] and one of them is actually being deployed in [a number of Lucent products], which is quite rewarding.

I think we'll come back to that point later. Recently, as Vice President, you've been the head of a group of a hundred researchers. How has that changed your perspective on things?

Actually, not a hundred, but sixty researchers. The major change [when you lead a large group like that] is that you have much less time to do research. So, you're moving from dealing with day to day research and reading papers and writing papers, to managing people, which is quite different from what I used to do in the university.

So, is [the change] a good thing or a bad thing?

Depends on the day. There are days that I am happy with what I am doing; there are a lot of days when I miss the ability to sit in my office quietly without interrupts and without action items from my boss.

It sounds like you need a sabbatical system at Lucent.

What I need to do is probably go back to academia.

Some of your projects at Lucent have produced publications but not had a business impact. How do you get research results into the development projects that will make it into the marketplace, for example, DataBlitz?

> [W]hen I hire people I ask them to produce results that are either going to add a lot of value to the company or will give them tenure at a top ten university.

When I am running the [Information Sciences Research] Center, what I want to do is have a good portfolio, a balanced portfolio, of things that are long term research and short term research. Some of the research [that we do at Lucent] we know in advance would not be of any value to Lucent[,] because it is [only] academically interesting. Some of the [research that we do at Lucent] had industrial potential but not within the domain where Lucent is, which is telecom; some of the [research that we did at Lucent] actually moved into startups that Lucent was sponsoring. So, the key is to have a balanced portfolio of long term, short term, and medium term research.

I see. So for the [research] that's short term, how do you get it into the products?

In industrial labs like Bell Labs, it's very, very tough [to move research results into products] because the business people are suspicious of researchers, and [similarly,] the researchers are [somehow] suspicious of business people. What you need to do is have [a research result] that the business unit needs, and you sit and work with the technical people [in the business unit] sometimes for years, convincing them that what [you] have is something that is going to change their product, and will enhance it, and finally if it works [out], [the business people] take [what you are offering] gladly.

Does it ever happen the other way around, where [the business people] come to your door wishing that you would do such and such for them?

The major problem that we have [with working with academics] is the issue of IP, and by IP I don't mean [Internet Protocol.]

Sometimes they come and ask you for consultations. They say, "I have a problem in X, can you help me with it?" Most of the time [when] you help them, it may be a couple of weeks help; and sometimes it actually turns into something that's more substantial---some [of the consultations eventually] turn into a product.

Do you think that that sort of consulting mode of research could be more widely deployed [in other] companies, or even within Bell Labs?

I had a rule that I [would not] have more than two to three PhD students a year […] [I]f you have ten, you don't have enough time [to] train them well.

I think that actually most industrial research labs are now in [this] mode, where if a business unit has some problem, they may find the appropriate person [in the research lab] and talk to [them]. When I was a professor in the university, I consulted a lot [for] other companies, and I'm doing the same thing now within Lucent. Instead of [the business people] going to the university and looking for appropriate people, they come to the internal [research] people, who help them with their problems.

How has Bell Lab's shift to more applied work in the last few years affected you and your group?

It's true that Bell Labs, in particular within Lucent

in the last few years, there's more [of a] shift to do something that brings value to the company. And bringing value to the company means two different things at Bell Labs. One of them is to do something that actually can be used in a product. And the other one is to enhance the brand name. So if you do long term research that is not of value to Lucent in terms of a product, it had better be something that gives you very good publicity. So if you can get a story in *Science* or *The New York Times*, or if you are on CNN and they are talking about [your research results], that enhances the brand of Bell Labs and that is very much appreciated. […]

How do you run your group at a time when fewer and fewer research organizations take the idea of academic excellence seriously?

I really have not changed the way I am running the group. I still believe in hiring the very best people; I still believe in a good mixture of long term research and short term research; and as long as I bring the right value to the company they let me do what I've done before.

How do you hire and nurture people who would be tenurable at a top ten university?

I'm not sure that there is any secret there. I just now unfortunately lost one of the people I hired about four years ago. [His] name is Mike Reiter, and he has become a full professor at Carnegie Mellon. […] You train [the people you hire] to do the right thing; when I hire people I ask them to produce results that are either going to add a lot of value to the company or will give them tenure at a top ten university.

So, you put it out front as an expectation.

One of the nice things about being in an industrial lab is that if I don't know the answer I can always pretend that there's an IP issue.

I put it out front: my expectation is that if you do academic work that is not of direct value to the company, I tell them it has to be the case that you will get tenure in the top ten. And if not, better do something that is of real direct value to the company.

How has startup fever affected your group? Has it been a good thing, or a bad thing?

Well, the good thing is that there is no more fever because there are no more startups.

It affected me in two ways. Probably the most immediate one was that it was much more difficult to hire people. Every time we interviewed someone, most of the time I lost them to startups. The other thing is in terms of retention; we lost some very good people to startups. There was not much we can do in terms of keeping [such people] around because the temptation of becoming very rich in a very short period of time was something I couldn't counter.

Lucent interviews many new PhD graduates from the top database schools. Are we on track in terms of the students we are producing from our schools?

That's a loaded question. Those people that I hire in databases are usually very well trained [in areas other than] databases. They have a very good broad understanding of computer science. These [indications of broad understanding] are the qualities that I'm looking for when I hire

people. Not only in databases, but any computer scientist that we hire, I'm looking for someone who has a very broad understanding of computing and algorithms. Many, many of these people who come in, within a short period of time, actually go through a transition where they start working on areas that are outside the traditional realm of databases. They may be working on networking, they may start working on real time databases, or start working on network management, where they draw on their strength from understanding databases. So, many of my database folks continue publishing in the traditional database [literature], but they are really moving into nontraditional databases.

I think that what we need to do is actually look for the right application domain for data mining. We actually use data mining in Lucent for a variety of projects.

Would you say that the candidates you see are better, [or] worse, than they used to be?

The quality of database graduates is probably the same [as it used to be,] when they come from the top universities. They are very well trained and I'm very happy with what I see in [those new graduates].

You've received some big awards for education. What are the secrets to being a great educator?

Be lucky and have very good graduate students.

I got the award, as far as I know, for two things. One is producing very good PhD students, and that's a question [of] being lucky [enough] to be surrounded by very good people; and the other is for the two books I've written that are used all over the world. So, [the secret to being a great educator is to] be a good writer and attract good students.

I can't believe it's just luck because overall, your former PhD students have done very well for themselves after graduation, but there are very few of them compared to most people of your seniority. Isn't it hard to have a big impact if you don't have a lot of students?

I don't know what "a lot" is. I think I probably had about 14 …

[I mean that you had very few] at a time.

I had a rule that I'm not going to have more than two to three PhD students [in the same] year. [The] reason is [that] if you have ten [students at once], you don't have enough time to spend with them and train them well and work with them. So I prefer to have fewer students but of very high caliber and I think my algorithm worked fairly well.

How is it that a professor with just a few students would want to manage 60 researchers?

Oh, the kind of work is totally different. When you work with the PhD students, [you] work with them much closer, you work with them hand in hand, and you really train them. When you have 60 PhDs working under you, [your] task is totally different. It was to come up with a vision [and] then let them run with it, and not be involved in the day to day research as I was with my PhD students.

When you say to have a vision and let them run with it---I've heard some people say that managing scientists is like herding cats. Have you found it to be so?

Sometimes, but if you hire the right people---and after all, I have four directors under me. Between four directors and myself, I think we manage fairly well to make sure that we have very high quality work done in research.

How should people in academia go about getting industrial collaboration?

That's a toughie. Having been in academia and consulted with industrial labs, [and with] industry in general, I know how to look at it from both sides. When I'm looking for collaboration from the outside, I'm looking for people who can add [real] value to my organization. […]

But that' s hiring them as a consultant. Do you ever do a joint research project?

Not really. Except that we have a lot of faculty who actually come and spend either a month or two months or even a sabbatical with us. We in general don't go and write contracts with faculty members at universities, and if [we do] so it's only for a very short period of time.

So my suggestion is, don't look inside the box. Look outside the box and start talking to colleagues that could use your expertise in databases in their own domain.

In the case of Lucent, then, the answer to the [original] question is that people in academia aren't going to have direct industrial collaboration [with Lucent]. They may come there for a visit, a sabbatical. You [the professor] join them, so to speak, if you want to collaborate with them.

It's very unlikely that I'm going to give some money to a faculty member at a university in the same way that NSF would give them money.

Well, you can collaborate [with someone] without giving them money. You can have a joint project without giving them money. But maybe that's hard in a corporate environment.

We have some joint projects. The major problem that we have is the issue of IP, and by IP I don't mean…

Internet Protocol.

I mean *intellectual property*. Lucent wants to make sure that if we work with a particular person at a university, the intellectual property stays with Lucent. And the only way we know how to do it is to actually sign a contract with that particular person, where that particular person agrees that for X amount of money, which is by the way very little, the intellectual property goes back to Lucent.

You are serving on a national committee for biodiversity. How did a computer science guy end up on such a committee?

I am actually no longer serving because the committee's job is finished. [In studying] biodiversity, one of the things they need to worry about is informatics. They have to worry about how to store information, how to retrieve it, how to catalog it, and so on, and [they deal] with very large databases. I was chosen to be the informatics guru. [The goal] is to build the database that will store the information that you need [for studying] biodiversity and ecosystems.

How did the committee [turn out]? Did you have recommendations?

That was a very interesting committee because we had people from all walks of life. We had Nobel prize winners from physics and biology. We had people from forestry and so on. We met for about a year and a half, almost every two, three weeks, in Washington DC, next to the White House, in the Executive [Office Building]. We generated a report that was presumably read by the President and by [Vice President] Al Gore, and we made a recommendation of how much money to appropriate for that particular activity. And that's where our task ended. So there was a big report with a very specific set of recommendations. […]

Is the database aspect of data mining dead as a research area? Is it still going on at Lucent?

I don't think that data mining as an activity is dead. I think that what we need to do is actually look for the right application domain for data mining. We actually use data mining in Lucent for a variety of projects. One of them that comes to mind, for example, is for fraud detection. If you want to figure out if a [phone call is] fraudulent, then data mining techniques are extremely useful. We are [also] using data mining techniques for dealing with issues of network management, [and] for fault management. There probably [are some ideas floating] around for dealing with denial of service attacks in the network, where data mining techniques can actually give you a clue that something bad is about to happen to the network and you might as well do something before it actually occurs.

Are those products or research areas?

Those are research areas and some of them are actually finding their way slowly into products. In particular, in terms of fraud detection in telephony, there is actually a product for doing that.

What is the role of database concepts in the problem of managing networks? I understand that that is one of your big interests now.

I think that if I had more time today, I would maybe go back to school and study some other areas where I think that the kind of expertise I have could be used, [such as] microbiology …

Actually, it is not the traditional databases that we teach our graduate students or undergraduates that we are using [for network management].

It is a collection of techniques that we use in databases that slowly [are migrating] into network management. I'll give you an example. If you look at network elements today (and by network elements I mean switches, routers, whatever), they all have something called a MIB which is in some sense a database where you keep information about all the events that occur in the network element. There are a lot of techniques we can use from database technology [to help make] MIBs much more up to date, much more modern than what are traditionally used in network elements.

And that's the whole topic? Or just one example?

That's one example. There are lots of other things, [but] now is not the right time to talk about it.

Oh, secret stuff!

It's not secret. One of the nice things about being in an industrial lab is that if I don't know the answer I can always pretend that there's an IP issue.

How is it that a professor with just a few students would want to manage 60 researchers?

I see. Well, we wouldn't want to give away any list of secrets here!

Compared to ten years ago, are the database papers being published now as important and of as high quality?

It's a difficult question to answer. The papers are definitely of the same quality that we had before. Whether they address the same caliber of issues that we had to deal with ten [or] fifteen years ago is not clear. If you look [back] over the years, we are basically [recycling and reinventing] the same algorithm[s] they've been using before for other domains. [For example], the same techniques that we used for the network model have been extended to the relational model, and then to the object-oriented model, and now to the XML model. So in some sense we are repeating things for different domains. [But] the quality of work is very good.

Is the repetition bad? Or [re]cycling?

It is the nature of databases and computer science as a whole. We are really driven by the technology rather than by anything else. When a new technology comes along, problems that were used before can be [addressed] now here. So if you look at XML, which is the hottest thing today, people worry about transaction management, they are asking themselves how you do optimization, they are asking the question [of] how you do views on top of XML, so it's the same thing we have done for the relational model fifteen [or] twenty years ago.

Do you have any words of advice for fledgling or midcareer database researchers or practitioners?

You mean like me?

Well… ok, do you have any words of advice for yourself?

Actually I don't have advice for myself, and if I did I'm not going to tell you. But I think [in] the database world, we solved a lot of the interesting problems. We need now to start expanding

beyond the traditional role of databases. There are some wonderful opportunities for working with scientists in other areas where our technology could be applied. So [for example], many years ago some people started working on databases for the human genome project. Unfortunately, we really didn't have the impact I was hoping for in terms of databases for the human genome [project]. But I think that there are a lot of other disciplines in biology and so on, where databases should actually be used. So I advise my colleagues in databases, look beyond databases and start talking to your colleagues in other areas.

I still believe in hiring the very best people; I still believe in a good mixture of long term research and short term research; and as long as I bring the right value to the company they let me do what I've done before.

So should they be looking at science, or are you thinking of things that aren't part of science too?

I'm basically thinking about science, for example bioinformatics, or …

Well, the catch I see there, at least for industry, is that it is very hard to make a lot of money selling databases to scientists.

I didn't think we were talking about industry; I'm talking about academia, database academicians.

[The question asked for advice for] researchers or practitioners, so they could have been anywhere.

I think, if you look at database work, let's say, in Lucent, or Bell Labs, we work with scientists; we work also with business people where database issues arise and you need to work with them. So my suggestion is, don't look inside the box. Look outside the box and start talking to colleagues that could use your expertise in databases in their own domain.

If you magically had enough extra time to do one extra thing at work that you are not doing now, what would it be?

I can think of lots of things, but I don't want to tell them in front of the camera. … I think that if I had more time today, I would maybe go back to school and study some other areas where I think [the] kind of expertise I have could be used. For example, [I might think of] going and spending a day or two [a week] studying microbiology again. So, go back to school and study something new.

If you could change one thing about yourself as a computer science researcher, what would it be?

Nothing.

Nothing! [Talking directly to the camera.] Ok, there we have it, folks. [Talking to Avi, who had been interrupted.] No, go ahead.

Nothing. [I'd change] nothing.

All right! We're all set for today, thanks.