An Interview with

JIM GRAY

OH 353

Conducted by Philip L. Frana

on

3 January 2002

San Francisco, California

Charles Babbage Institute
Center for the History of Information Processing
University of Minnesota, Minneapolis

Jim Gray Interview

3 January 2002

Oral History 353

Abstract

Gray discusses his childhood in Rome and education at the University of California, Berkeley. He explains the influence of Sputnik, Norbert Wiener's view of cybernetics and society, the social impact of computing, and the artificial intelligence papers of Newell and Simon in the shaping of his career. Gray describes his co-op position at General Dynamics, as well as positions with Bell Labs (Murray Hill) and IBM Research in Yorktown Heights and San Jose. Gray also describes his evaluations of computer models stimulated by the system dynamics approach pioneered by Jay Forrester, his brief role as a UNESCO technical expert in Romania, and his introduction to relational database design. The interview includes comments on computer privacy and research laboratory culture at International Business Machines, Tandem Computers, and Microsoft.

This is an oral history with Jim Gray at Microsoft Research in San Francisco. The date is January 3, 2002. This oral history is conducted for the CBI Software History Project.

Frana: Thank you for agreeing to sit for this oral history, Jim.

Gray: It is a pleasure to be here.

Frana: Thanks. I want to ask you first about your early life. Did you grow up here in the Bay Area?

Gray: Yes. I was born in San Francisco on January 12, 1944. My dad was in the U.S. Army, and in fact the war was still raging at the time. I have a sister who is two years older than I am, by the name of Gail. So we were here until the end of the European theater of action, at which point my father was assigned to an embassy position in Rome. The family lived in Rome for about the first three years of my life.

Frana: Do you recall any of that?

Gray: I recall some of it, images mostly. It was a very unusual life. We had a governess. Rome was in a mess. The economy was just absolutely terrible. Being hooked up with the American Embassy we lived in what I would call a palace, even today. There was lots and lots of

entertaining. As far as I can tell my father's job was espionage, and trying to figure out whom he could trust and whom he couldn't trust. The communists were very influential and we were trying to minimize their influence.

Frana: He never divulged too much information to you about this?

Gray: He did later, yes. Certainly at the time I was clueless. I was two or three years old. What I really saw was just parties going on all around me. It seemed like an endless stream of visitors and parties.

Frana: And your mother?

Gray: My mother. Yes, she was close to us. She was born in San Francisco as well. She had gotten a degree in education from San Francisco State and had taught here. I guess they were married in the 1930s. He had been in the military reserve, and when the war started he got called up. She continued teaching and had two kids. When the war ended she put her career on hold and moved around with him for a while.

Frana: Did you move around a lot? When did you move back?

Gray: The army is a pretty mobile organization. People typically do four-year tours of duty. We moved from Rome to Virginia and spent about four years there. At that point my parents divorced and I moved to San Francisco and lived here for most of the rest of my life.

Frana: Were any of your family in any way remotely interested in mathematics?

Gray: Not so much mathematics as technology, I would say. My father was an amateur inventor. I remember him struggling with trying to make a refrigerator that automatically made ice cubes. He spent years at that—other people were more successful than he was. The one invention that ended up paying off very well for him was putting typewriter ribbons in plastic cases, and so having a cartridge typewriter ribbon. He patented that idea and it generated a huge royalty stream for him. It probably paid for a lot of his extracurricular activities after he got out of the army.

[pause]

Frana: How did you made your decision to go to Berkeley? You studied mathematics?

Gray: I did. I showed up at Berkeley about 1961 as a freshman. It was pretty easy to make the decision to go to Berkeley, in the sense that it was essentially free. It was $67 a semester to go to Berkeley at the time. We were not terribly well off. I had applied to the Air Force Academy as well, and I didn't get in.

Frana: Eyesight?

Gray: No. It is actually very competitive to get in. You have to be recommended by a Congressman. It is also free and they actually give you a very good education. It is a little

regimented. I was not enthusiastic about the regimentation, you know, this is pre-Vietnam. The military has gone through a period of being unfashionable and that might change given recent events. I ended up going to Berkeley because: (a) it was affordable and it was a very good school, and (b) they let me in. So it was a pretty good combination. My sister was also going to school so it was pretty difficult for my mother to afford sending us both to college. We had to think pretty carefully about the economics of it.

Frana: Your father's money didn't—

Gray: He was obliged to pay alimony, but he wasn't obliged to send kids to college. He had his own demons to fight. He was not very involved in our upbringing.

Frana: At what point did you hear about these miraculous things called computers, which led you to enroll in that numerical analysis course?

Gray: Well, I think actually in high school I had heard of and read some of Norbert Wiener's books. He had a paperback that came out entitled *Cybernetics and Society*. It was difficult not to hear about computers after Sputnik had gone up. Sputnik had a huge impact; it is difficult for people to appreciate this. I would compare September 11th to Sputnik. It just had this galvanizing effect on people.

Frana: Your friend Don Chamberlin has said much the same about Sputnik.

Gray: Yes. So you know technology had a fairly positive connotation. People were concerned that it was going to eliminate jobs and therefore create unemployment, but that was about the only thing that people could see as being a downside. They weren't concerned about alienation so much. So I think my first connection to the idea of information processing was the work of Wiener. I was also very interested in mathematics. I found it a good way of thinking about the representation of knowledge, and I thought that was one of the fundamental problems that we face as philosophers. Incidentally, I started out a physics major, not in the engineering school. Like all freshmen, you actually go into the School of Letters and Science or School of Engineering, and when you have been there for two years you declare a major. But somehow, in the back of your mind, you say, 'Well, I am going to go into physical sciences, or I am going to go into pre-med, or I am going to go into history.' So you have to have something in the back of your mind telling you where you are headed. I was in the School of Letters and Science and—like most freshmen—I changed my major once a week: 'Oh, here's some neat idea,' I remember thinking often. 'That sounds good.' For a while I thought I was going to go into philosophy. In fact, I took a philosophy class, but it was clear that the way they were approaching representation of knowledge was not going to scale to the kinds of problems that interested me. I mean, they could represent the simple predicate calculus but they couldn't represent more complicated concepts. It looked like computers actually offered an opportunity to represent things in a much better way. Then I heard about the Newell and Simon proving theorems, and I thought that was just fascinating.

Frana: The Logic Theorist and General Problem Solver?

Gray: Yes. And so those were the sorts of seeds that had been planted. I went and worked as a co-op student because I was basically working my way through college. My first job was working at General Dynamics in San Diego. My job was as a computer. I had a Monroe calculator and the guy would give me tables of numbers. I would grind out more tables of numbers. One of the other co-op students was using this thing called Fortran. He explained to me that what I was doing could be done very easily with Fortran. I didn't actually follow up on that, but it seemed like that was definitely going to be the way things would happen in the future.

Frana: Since you'd read Wiener's book did you consider the computer a way to solve social problems?

Gray: Well, that was certainly Wiener's main focus. You have to appreciate that I was seventeen or eighteen years old, and that was an interesting concept. It seemed very far from reality. It seemed more like a vision. Later, when I was a graduate student I came back to that problem when the Club of Rome was publishing things about a basically a dystopian future.

Frana: In the 1970s I carried that book around and worried about the world myself.

Gray: Yes, right. I looked into *The Limits to Growth* and built those kinds of models. I tried to understand what they meant and whether they could be used for decision-making. And the answer is: models like that could be used for decision-making, but those models were far too simplistic to make reliable predictions. The fundamental thing that those models did not have is substitutability. You couldn't build skyscrapers in that model because there was 'land,' and land

was defined as all one-storey buildings. You also couldn't substitute one form of energy for another form of energy. But we are getting way ahead of things.

Frana: Yes, this was when you were at IBM.

Gray: Yes. I did that work at IBM, but it also really started at Berkeley.

Frana: You had heard of Forrester's 'system dynamics' at that point?

Gray: In fact, what sort of started it was reading Forrester's *Industrial Dynamics,* and then it was called *World Dynamics,* and then there was *The Limits to Growth.*

Frana: Which was actually Donella Meadows, right?

Gray: Meadows and The Club of Rome. Right.

Frana: Sorry. Let's continue with the chronology.

Gray: So I was a freshman and had gone off to co-op. Actually, I had done okay as a student in my freshman year. I did very well in some classes and mediocre in others. Chemistry was my graveyard. It was a five-unit class and I got two C's in it. What it did to the grade point was just hellish. I always seemed to be a week behind. I don't know if you have ever been in a class like that?

Frana: Organic Chemistry—I found it bewildering.

Gray: Yes. I wasn't sure I wanted to be in the university anymore. So I went off and worked in industry for six months. It became crystal clear to me why I wanted to be in the university: it was a lot more fun. The people I hung out with and the work I was doing in industry were just dreadful. I was interested in not spending the rest of my life doing that, and not being like the people I was working with. I don't mean to denigrate them in any way. It just wasn't what I aspired to. I went back to school with much more resolve. And all of the sudden, my grades were much better, and I was much more focused.

Frana: When was this?

Gray: This would have been 1962 or 1963. I was a co-op one more time. After I returned from that, I started taking courses in computing because I was really getting more and more excited about it. I took a numerical analysis course. I also started taking some graduate classes in electrical engineering. All of the computing classes were graduate classes at that time. It was hard to get into these classes. There were prerequisites, and I hadn't taken the prerequisites. I had to go talk to the professor, and he said, 'Well, you are welcome to take the class, and if it is too hard you can always drop it.'

Frana: What kinds of courses were these?

Gray: What was being taught was finite state machines. ‘Automata theory’ is what it would be called now.

Frana: Discrete systems?

Gray: That’s right—and recursive function theory, too. Discrete mathematics is what it really was. I had been supplementing my income by reading math papers, and I was actually pretty good at this math stuff. The discrete mathematics that they were doing in the engineering school was pretty easy compared to the discrete mathematics that they were doing in the Math Department. I was among the best students they had, although I was still an undergraduate. If you are a faculty member, and you see a bright undergraduate, this is a very good sign. You try to grab those people because they are full of energy and don’t know they are supposed to have a life. Mike Harrison offered me a research assistantship. This was way better than being a Reader in the Math Department. The pay was better and the work was more interesting. That is how I met Mike. I was an RA the latter part of my junior year and my senior year. During that time I wrote two research reports. At that point, I was still a pretty mixed up kid. I had married, and we wanted to travel around and see the world. I figured I had all the time in the world. I still think I have a lot of time. I wasn’t in any hurry to start a career. So my wife and I moved to New Jersey, which is where she was from. She got a job teaching. I got a job at Bell Labs working on a digital simulation, which was going to be part of Multics. Bell Labs, MIT, and General Electric were putting together this great time-sharing system. Loretta got a job teaching and I got a job working at Bell Labs. Bell Labs was wonderful about supporting continuing education. Two

days a week I would go down to the Courant Institute in New York and take classes. Basically I worked three days a week and went to university two days a week.

Frana: Had you graduated from Berkeley at this point?

Gray: I had a bachelor's degree and was about to get a Master's. So I was enrolled as a Master's student at the Courant Institute. I was taking really great classes from wonderful professors, and then going to the Museum of Modern Art for the afternoon. It was a really good life, in fact.

Frana: Did you get to meet Fernando Corbató?

Gray: No, I didn't. I was a lowly grunt in the Bell wing of this thing. Frankly, the Multics project as a whole was going very badly. The project was late, and I was actually using the Dartmouth timesharing system to do development on it. The whole goal of this was to work for a year to make enough money so that we could take off five years and go travel around the world. This was early in the Vietnam War—something like 1966. People from Berkeley were fundamentally alienated from the rest of the United States, and they were especially alienated because of the Vietnam War. There was a drug subculture at Berkeley. I mean, it was weird. On the other hand, it was unpatriotic to be against the war at AT&T. We were eager to get out of New Jersey by the end of the year. We made some friends, but I never really felt like I wanted to be there forever. I tried very hard to blend in at AT&T. I wore tweed jackets and so on. But I did drive to work on a Triumph motorcycle. It was pretty clear that I was not the ordinary employee, although there were others at Bell Labs like me. I had gone to college with Ken Thompson. He was not that

dissimilar. It was not like there was a generation gap at AT&T. The younger generation—many of them went through fairly ordinary experiences. Some of us had gone through fairly disorienting experiences.

Frana: Was there a split between East Coast and West Coast culture in information technology?

Gray: Well, AT&T didn't really have a West Coast. It had some things in Indian Hill, which is outside of Chicago, and it had Murray Hill. There were a bunch of places in New Jersey, and one place out by Chicago. It was really Bell Labs against AT&T. If there was any culture shear it was between the people who had gone to college and the people who had not. That is the simple way to describe it. It didn't seem to me to be generational.

Frana: So did you go travel the world?

Gray: We started. We came back to Berkeley and spent about three months here. We had lots and lots of friends at Berkeley and it seemed like a paradise given the weather, the intellectual climate, and friends. And so yes, we started traveling. We were probably on the road for about two months when both Loretta and I decided that we didn't actually like traveling as much as we liked just being at Berkeley and being around our friends. Traveling sounds really romantic and great, but it is actually a very lonely life. Every day you meet new people, which is exciting, but then you have more or less the same conversation the next day with some new people. You don't form long-lasting friendships. If you are a couple, you have each other. I think there are people for whom that is just the perfect life, but we weren't those people.

Frana: Did you complete your Master's at Bell Labs?

Gray: No, I didn't. I took four classes at NYU, and had a wonderful time, and wrapped up the project I had been working on. My wife's teaching responsibilities ended. We basically spent a year there and came back to Berkeley. So, as I said, I didn't really enjoy traveling all that much. I came back and returned to graduate school. I was a graduate student in the newly formed Computer Science Department at Berkeley.

Frana: Working again with Michael Harrison?

Gray: Yes. Mike Harrison was my advisor. It was an interesting group of people. Steve Cook, Dick Karp, Butler Lampson, and Jim Morris were on the faculty. The people at Stanford were pretty astonishing, and we were on pretty good terms with them. We had reasonably good contacts with MIT and UCLA. It was a very small world. Something, I guess, that is difficult to grasp these days. Everybody knew everybody. One had even a stronger feeling of that back in the 1950s. They had these conferences where quite literally everybody in the field showed up at the conference. Even in the 1960s Mike Harrison knew lots and lots of the principals. Through him I met many of the principals in the field. Mike was quite close to Seymour Ginsburg who was one of the leaders in the field in formal language theory. I got to take classes from Dick Karp, Steve Cook, Mike Harrison, and Butler Lampson. It was really a great education.

Frana: Was Mario Schkolnick there?

Gray: Mario was a graduate student there at the time, yes. Peter Deutsch was a student there. So was Charles Simonyi—and Jean LeClerc, who is a big wheel in France these days, and Herve Gallaire who is a CTO at Xerox.

Frana: Your Ph.D. was in programming languages?

Gray: That's right.

Frana: How did that happen?

Gray: Well, I was really impressed with how elegant BNF was. There is a duality between the grammar model, and the generative model, and the machine recognizer model. You could have Markov languages, or generative languages, or post systems, that are equivalent to Turing machines. And there is a computational complexity hierarchy that says, 'Context free grammars are equivalent to push-down automata, and general rewrite rules are equivalent to Turing machines, and regular expressions are equivalent to finite automata.' Harrison had been operating primarily at the bottom of that hierarchy either with finite automata or with pushdown automata, and either with regular expressions or context free grammars. And the theory of parsing, which is to say, given a language definition—BNF or context free grammar—and given a particular sentence, how do you parse that sentence with that grammar? And in particular, how do you generate a parser that is efficient? That was an area that was still not completely understood. Jay Earley had done some very good work and he was on the faculty there. Harrison

had been working in that area too. In particular there was a way of doing parsers called precedence parsing, where you take and treat certain of the tokens as operators and certain of the tokens as operands, and you assign precedence to each of the operators. No one had really characterized what kind of power that kind of parser had. So my thesis was primarily characterizing the mathematical power, or the parsing power, of operator precedence grammars. That actually gave rise to a need to be able to talk about the equivalence to grammars—that this grammar is just as rich as this one. If you had a grammar that just had phrases and didn't understand the difference between noun phrases and verb phrases, that would be a simplification of grammar that did understand noun phrases and verb phrases. So the second part of the thesis was trying to have a taxonomy that says that one grammar is essentially a superset of another, and that there is an effect on the homomorphism from one to the other. And they are equivalent, of course, if there is an isomorphism.

Frana: Behavioral theories—were you looking at those as well?

Gray: Well, there was a group, and I was really quite interested in it, that was started by Pravin Varaiya. I think Gene Wong and Pat Mantey were also members of it. We were looking at the social applications of computers. Clearly computers were a big help for industry. We wanted to know what they were going to do to help government and society. We did things like read Jay Forrester's *Industrial Dynamics* and *Urban Dynamics*. Pravin was a control theorist and a system theorist. He was trying to take a systems theory approach to modeling urban problems. His seminar grew into some early work that I did, basically reproducing what Forrester had done. It was difficult to reproduce Forrester's results even though there was this book, and he published

his code. You had to have the simulation software Dynamo**,** and you had to have this kind of computer system, and so on. So I basically built this Dynamo simulator from scratch in SNOBOL and then did the simulations again. That turned out to be a very useful tool for us to have, and allowed us to do experiments that others couldn't do. We had pretty good computing resources at Berkeley because we had the Project Genie Timesharing System. I had gotten what would be called very surprising results. It turned out that with the model, if you looked at certain parameters, the unemployment rate was like 75 percent. We would say, 'You know this doesn't seem like it is a very good model of what is going on.' Parameters were very, very sensitive and they were sensitive in strange ways. There was a great interest in applying computing to social problems. This was in part due to the social consciousness of people in Berkeley. We thought that we ought to be doing something for society directly.

Frana: You worked on this problem during your two-year post-doc at IBM?

Gray: I did that during the two-year post-doc, and during my graduate work. This work started before I actually finished my dissertation. Relatively few people were making contributions in this area, and so the modest contributions that we were making actually ranked fairly high on a national scale. The people at IBM had been hearing about the Club of Rome**,** and *The Limits to Growth*, and so on. The Club of Rome had a very dystopian view of the way things were going to go. People in IBM Research were being asked whether they had any comments to make on this, when I showed up on IBM's doorstep. I actually had been working on a capability-based operating system called Cal TSS, which had been started by Lampson and Simonyi and Deutsch before they went off to the Berkeley Computer Corporation. I think if you talk to Butler you'd

see a dovetail between these two stories. I inherited this project that the National Science Foundation had funded to build a timesharing system on the CDC machine that was at Berkeley. They had two CDC machines, and one of them was used to develop timesharing. Money had been raised and they needed a titular head. I had the credentials, which is to say, I had a Ph.D. and was a researcher.

Frana: Was this the 6600 machine?

Gray: They were. Yes. It was a really nice design. There was a guy there named Howard Sturgis who went on to work in Xerox PARC, as well as Paul McJones, Dave Redell, Vance Vaughan, and myself. Jim Morris contributed some code to the project. It was a stellar group. These were very, very talented people, full of energy, with great enthusiasm, and an interesting design. In the end it did not have very much direct impact, but certainly it had a huge impact on the people who worked on it. So that's how I finished my Ph.D.

Frana: In 1969?

Gray: In 1969, that's right. I loved being at Berkeley, and I wasn't very excited about the prospect of going elsewhere. IBM had these post-docs, and I got one to be at Berkeley for two more years. I had basically no responsibilities. It was a great situation: I was doing productive research and enjoying life enormously. I had a daughter almost the same time the thesis got finished. Unfortunately, my marriage came to an end at about that point, or a few years later. I think it is not uncommon for this to happen. People go through graduate school, they get out on

the other end, and they are different people than they were when they got married in the first place. So we separated. I went to work for IBM in Yorktown Heights. I needed the job. At that point all the jobs were on the East Coast. I thought of going to academe, and I am not exactly sure why I went to industry rather than academe. I worked in the General Sciences Group at Yorktown, ostensibly continuing the simulation work I had been doing. I tried to make connections with people in New York City because they definitely had urban problems. I ended up working primarily on what is called 'cadastral mapping,' which is land-use mapping. I spent a great deal of time trying to computerize the assessor's records. But the technology wasn't right at that point. It was too hard to digitize things. But in that era I met John Cocke. I think the mentors and role models we meet influence all of us, and John was definitely a role model for many of us. John got me excited about the idea of scaleable computing. He said, 'Wouldn't it be great if you came up with an architecture where if they just gave you twice as much money you could do the work twice as fast?' We don't actually know how to do that. Within bounds, Beowulf clusters do that, but not all problems fit on a Beowulf.

Frana: So early in 1971 you got interested in this?

Gray: Yes, I would say it was 1971. Cocke had always been interested in scalability problems. Mostly he had been working scale up—which is build bigger and bigger and bigger machines—but it was clear to him that this had its limitations. At a certain point you do have to scale out.

Frana: He was at Yorktown Heights?

Gray: He was at Yorktown Heights, as were Don Chamberlin, and Frank King, Ray Boyce, and Peter de Jong. I met lots and lots of people. I was a bachelor and Westchester County is pretty quiet. I had nothing else to do except work.

Frana: Leonard Liu was also there?

Gray: Leonard Liu was there.

Frana: Was he your manager?

Gray: No, he wasn't. I was actually in the General Sciences group. But at that point I was all over the map. I was doing theoretical computer science. I was doing operating systems work. I was doing simulation. I was working in urban modeling. I was doing interesting things in each of those areas, but not with very much depth. I needed a little focus. I arrived approximately Halloween eve and I remember going to a really great Halloween party in Westchester County. There were really interesting people there, but it was the onset of winter. I was alone, and I was going through a lot of emotional problems. This was when I was challenged by winter. I went to my manager in February and said, 'Dinos, I am going to be okay. I am going to get through this winter, but I am never going to spend another winter on the East Coast again in my life. I would like to apply for a transfer to California.' And Dinos said, 'Sure.' So June came and I said to him, 'Okay, well, I am leaving.' He said, 'I thought you were kidding.' I said, 'No.' I had arranged to go work for UNESCO for the summer in Romania as a technical expert. This was again based on my desire to do social service.

Frana: You were a visiting computer science professor then?

Gray: I was a UNESCO expert at the university for a semester. UNESCO experts go to these less developed countries and are paid more than IBM employees and researchers for a lot less work. I was being paid way more than anybody I was working with and that creates a certain amount of tension because they could really use the money for other things. We really didn't have a role besides teaching. In other words, we could have given people advice, but there was really no mechanism for us to give advice. UNESCO was not particularly concerned about this. They were in Paris and more concerned about the relative value of the dollar versus the franc and things like that. I went with certain motives, which were rather unusual for the other UNESCO experts. This was a summer vacation for most of the other people doing this work.

Frana: The others easily accepted the situation.

Gray: They accepted the situation. That's right.

Frana: What happened?

Gray: I was there for three months and did the best I could. I really wished that I could have been more helpful than I was. An amusing story: I had read most of Marx, and I thought that communism was actually a pretty good idea, and that capitalism was a pretty bad idea. And so it is quite possible that when I went to Romania I was the only communist in Romania. After about

a week or two there were no communists left in Romania, because I got to see first hand what it really meant when things were nationalized and everything belonged to everybody. In a kibbutz I am sure it works great. But we have not figured out how to scale communism beyond the family or the kibbutz. I mean, even the Israelis have their problems. They don't have a national kibbutz.

Frana: Very localized.

Gray: Yes, very localized. And in fact very specialized people. Certainly in this period there were communes in California and New Mexico. I visited them and they fundamentally operated on a communist model. Most of them at this point have evaporated, twenty years later. It is very hard to sustain a commune, especially in our culture.

Frana: Human selfishness gets in the way?

Gray: Well, society needs to have people who are motivated, and there just isn't enough motivation. If you have two organizations, one that is laid back and the other that is highly motivated, there is going to be a sort of Darwinian interaction between the two of them. So, I'd worry that a really communist state would not survive very well with non-communist neighbors, in the sense that their neighbors would probably tend to take over. This is not something I've thought very deeply about, but it was a lesson for me that Marx had done the intellectual experiment, but not the physical experiment. There are a lot of things like that. If you write a program, you say, 'This ought to work.' Then you run the program and discover there are many things that you have not figured out. That is just a trivial example of the problems you face when

you start experimenting with a system. Anyway, I came back to America and I came back to Berkeley and tried to put my marriage back together again, not successfully.

Frana: With Loretta?

Gray: Yes. We decided to get divorced. I looked around for a job. This was one of those horrible times where there were no jobs. You know, there is this cyclic thing in the economy. Every four years things would get good, and every four years things would get bad. People have again kind of forgotten this, but it used to be a serious problem. People would come out of the university at a certain point, and if you caught the wave right, people would bang on the doors and recruiters were everywhere. But if you got the wrong year, you'd come out and no one wanted to hire you. But at any rate, I went around and looked for jobs. There were not a lot of choices. I thought of going to a university. I wasn't in any great rush. But I knew the people at IBM. I'd worked at IBM, and I liked IBM as a company. I thought they had integrity, which was something I really worried about. I thought they were doing interesting things. I liked their products and people, the way they treated their customers, and their focus on honesty. I also liked the fact that they took care of their employees in bad times. They did not lay people off. So I thought it was a great company. I probably could have gone to work at Xerox PARC, and I don't know why I didn't really work that agenda very hard. I went and visited my friends at PARC, but I never actually thought about going to work there. So the long and short of it is, I ended up working at IBM in San Jose. I started out working on a project trying to build a database machine, which is to say a machine designed explicitly to run database applications. The theory was that if we microprogrammed the computer to do databases it would be better than if we did it with ordinary

software. I thought that it was a crazy idea, but it is good for researchers to work on crazy ideas and so I thought maybe something would come of it. We worked on it for about six months and concluded that we didn't see how by changing the level of abstraction downwards that we were going to make things better. It looked like things would be a lot worse. We actually wrote it in millicode rather than microcode. System/3 had a processor level abstraction, which is a fairly conventional processor except that it has a level store session. The designers had written a lot of what I'd call assembly language on top of it to make the System/38 or AS/400 architecture, which is in fact, a capability based machine. You might want to talk to George Radin or those guys about System A.

Frana: Did you contribute to System A when you were at Yorktown Heights?

Gray: Well, it was underway. I certainly talked to the people and was very much in sympathy with what they were doing. But then it turned into something called Future System, or FS.

Frana: Oh, yes.

Gray: I was one of the staunchest critics of FS. I said, 'This is too complicated. We can't actually build this and there are too many unsolved problems here. This is a great research project, but not a great product that you want to ship.'

Frana: Your Multics experience helped you see that?

Gray: Yes, and the Cal TSS experience. I mean these people had chutzpah. The System 38 has processes in millicode—that's fine. They want levels—that's fine. But they had lots and lots of abstractions like queues and events and so on, which are very hard to get straight. Anyway, so Ted Codd was wandering around saying, 'We ought to use relational calculus as the database query language.' There was growing interest and enthusiasm for Codd's ideas, especially in the research community. So Codd was allowed to start a little project. Dines Bjørner, Ted Codd, myself, and Irv Traiger were in this core group that was experimenting with a system called GAMMA. Ted called the system GAMMA. It was essentially a relational calculus. We were going to build a relational database system. There were other people around the world building relational database systems inspired largely by Ted. There were people in IBM at the Peterlee Science Center in England. There were some people at Berkeley who were beginning to think about these ideas. There were some people at MIT who built a system. There was a group at the Cambridge Science Center. Various people were working in this area. We were not alone in doing this.

Frana: It would be difficult to establish priority?

Gray: Absolutely. David Childs will tell you that he had this idea long before Ted. It is astonishing how many people will point to Ted and say, 'Ted is the guy that we got the idea from. When we read Ted's paper we finally understood it.' Now in fairness, there are people at IBM who say, 'Well, we gave Ted the idea.' The long and the short of it is that the research management at IBM was looking at what was going on in San Jose. They were getting hints from the outside that the mainline database research that was going on in San Jose was adequate, but

not going to be a breakthrough. There was a group of 'Young Turks' in Yorktown, known in San Jose as the Yorktown Mafia, who said, 'What those guys in San Jose are doing is linear progress, and what is wanted is paradigm shift. Here it is. Send us out there and we will save the world for you.' This is Leonard Liu, Frank King, Ray Boyce. Don Chamberlin and Mike Blasgen tagged along for the ride. Fundamentally, you talk to Don and he is a very, very, very nice guy. He is not an empire builder, but he had ideas, and other people were willing to take those ideas and shape them for him.

Frana: The System R oral history prepared by Paul McJones describes relational databases as 'a logical next step.' You see it as a paradigm shift?

Gray: Yes. What Charlie Bachman was trying to do, and what his heirs were trying to do, was to make it easier and easier to navigate through circles and arrows. They had sets, and they had memberships to sets, and so on, but they did not have a set oriented non-procedural query language. When you look at the Internet right now, it is really back to the Bachman world of circles and arrows. We are beginning to see hints of things, which make it easier to navigate, like Google. They go and they take the information and reorganize it into ways that facilitate associative access. Is it a paradigm shift? It certainly seemed to be so to the people at the time. There were two camps, and they were warring with each other. Each camp couldn't understand what the other was talking about. They had completely different assumptions about what was important.

Frana: Do you accept the argument that the navigational model was trying to achieve some sort of data independence?

Gray: Oh, yes. Data independence they very much wanted, but they were not looking for non-procedurality. The CODASYL-DBTG had schemas and subschemas, and subschemas were data independent. The main thing I think the relational model gave was an ability to state queries on order of magnitude more simply. I mean, just fewer lines of code. We thought that people would write SEQUEL. Well, forget that. The simple fact is that tools can now generate non-procedural queries and not have to know about what the user sees and what the physical structure is. And these non-procedural queries get executed. If we were still in the world of the DBTG it would be much, much harder to make the tools we have today, where people just get tables and say, 'Take this from here, and take this from here, and combine them.'

Frana: Right. Do you remember the Great Debate at SIGMOD 1975 between the network and relational people?

Gray: I wasn't there.

Frana: You weren't there?

Gray: No. I do remember Bachman. During this period there were lots of people who couldn't talk to each other. There were people who were really angry at each other. Fortunately, I was down in the plumbing. I just wanted to get the bytes to and from disk. I just wanted to run the

programs in parallel. I wasn't pissing anybody off. I was doing stuff that nobody else wanted to do. People were perfectly happy to talk to me about what they were doing. I was friends with the IMS guys, and I was friends with the DBTG guys, and I was friends with the relational guys. All this flak just sort of missed me. I was just doing the operating system level stuff.

Frana: Ninety percent of the Fortune 500 companies still have IMS.

Gray: Oh, sure. Yes, IMS is a big moneymaker for IBM. There are ways to go around it, but they don't need to go around it. It works fine for them. You know the Fed clears every night using IMS. What the Fed is doing is very, very, very standard. They are just clearing checks. The year 2000 came and they had to go to four digit dates instead of two digit dates. Things like that are a major change for them. The airline reservation systems are actually running on systems even more primitive than that. It is unfair to describe the IMS guys as Amish. They are very proud of what their systems do. Their systems are up, and very, very stable. One of the things that causes systems to be unreliable is change, and so the IMS systems are pretty static. They just sit there and year after year after year, they do what they do. You know, if you think about banking it really hasn't changed that much. Automatic clearinghouses work more or less the way they used to. The equipment is more modern and cheaper, but I suspect many of programs were written twenty or thirty years ago.

Frana: So who were the great populizers and the proselytizers for the relational model? Ted? Chris Date? Yourself?

Gray: You know, not so much. I was actually pushing this transaction stuff. I took all of the relational stuff for granted. Don Chamberlin wrote many papers. Mike Stonebraker and the people at Berkeley wrote papers. What happened is that the academic community found DBTG and IMS pretty complicated. It wasn't elegant and there wasn't a theory associated with it. You couldn't prove theorems about it, or at least they didn't figure out how to. And along came the relational model with query optimization, and transactions, and security. The data model was simple enough that you could state it and then start reasoning about it. This caused what used to be called the ACM Special Interest Group on File Description and Translation (SIGFIDET) to turn into the Special Interest Group on the Management of Data (SIGMOD). They changed their name and saw a complete change of cast of characters. There began to be academic departments, and those departments started producing students. Some of those students went off to form other departments, and some went to industry. It was a self-reinforcing system. Now there are people who are database professors at one or another university, and almost all of them are from the relational tradition. In academe I would say it was really Stonebraker, Wong, and Rowe at Berkeley who were leaders in this. Dave DeWitt and his students came a little bit later at Wisconsin. One hundred years from now IMS is not likely to be running. I suspect that the number of IMS systems is not growing right now. The number of licenses might be growing slightly because people are buying more computers to run the same applications, but in terms of new applications coming in line I suspect it is static. They lose some every year.

Frana: Before we move on—was it Leonard Liu who encouraged you to get into database rather than operating systems and networks?

Gray: Yes. But it's interesting. As I say, I was all over the map and I continue to be. To this day I am interested in more than one thing. I am very interested in almost everything technical. Leonard consciously came into my office and had a heart-to-heart talk with me. He said, 'You know, IBM has too many programming languages. IBM has too many operating systems. If you want to do work in programming languages or operating systems, the best thing you could do is kill off a couple of them. On the other hand, we've got big problems in networking, and we've got big problems in databases, so if you are looking for an area where we actually need new ideas, as opposed to consolidation, it is in the area of databases or data communications.' It was not a Machiavellian discussion. It was just an honest assessment on his part of where the excitement lay. He didn't tell me, 'Stop doing what you are doing and start doing something else.' It was just advice.

Frana: Did that come as a shock to you?

Gray: No. I thought it was eminently reasonable. I remember the conversation and it had exactly the impact that I think he desired. I was working for IBM, and they were paying my salary. I felt some sort of obligation to them to do work that was relevant to them, and would make them prosperous. As much as I am working at Microsoft now, and Microsoft pays my salary, and I try to do things that are helpful to Microsoft.

Frana: Now, transaction processing—that was something you were working on with the database machine? That is something that the DBTG committee struggled with.

Gray: The DBTG had a kind of concurrency control and a kind of recovery. The concurrency control was, as you can imagine, manual, because everything in DBGT was manual. You could get a lock on something and locks were exclusive. Only one person at a time could have a lock on something. When you were done you would say 'unlock.' There was really no transaction concept at all. IMS actually was a little bit fancier. While you could explicitly lock something, IMS would more or less acquire the locks for you. At the end of your work you could say, 'Sync. Sync point.' And 'sync' essentially was commit. You could also say, 'I wish I hadn't done this.' You can say, 'abort' or something like abort. IMS had evolved over time. It had started out very much like DBTG with nothing. It acquired the ability to declare resources in advance so the transaction would not run. You had to declare what locks you were going to need, and that did not work out very well. The thing I just described was something called the 'program isolation feature.' There was pretty good flow back and forth between the SEQUEL group and the IMS group. I knew the guys who were working on the IMS side. They were not very far away, up in Palo Alto, and I would go and visit them. Other people—Franco Putzolu, Frank King, and so on—would also travel back and forth. They viewed us as being kind of strange. We were from IBM Research and weren't doing anything 'useful.' But they were civil. In fact, when we talked to them about things that we were doing, they showed some interest and gradually began to think, 'Well maybe these guys could actually do something useful for us.' Incidentally, they viewed DBTG as the enemy and they viewed relational as the enemy. DBTG was not IMS. DBTG was CODASYL and the network data model. They had a hierarchical data model. IMS was face to face with products from Burroughs and Univac and others that did things in the DBTG way. IBM was under increasing pressure to have an implementation of this thing that was likely to become an ISO standard. The company was eager to do what they could to make DBTG

fail, frankly. We sent people to the DBTG committee but we did not want the DBTG to be successful. It wasn't going to be good for IBM. SEQUEL had the same status. Even though SEQUEL was something that grew out of IBM, IBM didn't have any SEQUEL products at the time. For SEQUEL to be successful, it meant that IMS would be less successful. And IMS was IBM's future direction. That's a whole other story. But so I talked to the people. And you know, transaction processing wasn't called that, but the germs of the ideas were in various systems that were floating around. I stepped back and said, 'What is going on here? What are the right concepts and what are the right definitions? And if you take those concepts and definitions, what are the implications?'

Frana: So in part you were trying to rationalize what was already out there?

Gray: Right.

Frana: This is where you felt the core properties—

Gray: Yes. The core properties are 'ACID': atomicity, consistency, isolation, and durability.

Frana: Right.

Gray: That is actually something that was dreamed up by Theo Härder and Andreas Reuter in a *Computing Surveys* article.

Frana: The acronym?

Gray: Yes, the acronym is theirs. The idea of atomicity was certainly in the transaction model. And the idea of durability—much of the early work in transactions had to do with concurrency control, which is just the atomicity issue in the isolation machine. And so the durability issue, which has to do with recovery, really was treated separately and has been until very recently. You will see a paper about recovery, or you will see a paper about concurrency, and very few papers talk about recovery and concurrency at the same time because the mechanisms are usually completely separate. One is locking and the other is logging, or multiple copies of things—redundancy for recovery. There are people who unify the two concepts. One of the most interesting things is that over the last ten years the DB community has in fact done a wonderful job of unifying keeping multiple copies of things to allow concurrent access, and keeping a journal for recovery, and keeping multiple copies for recovery.

[pause]

Frana: As we are running short of time, let me skip ahead to a few big picture questions: Has the research climate in databases changed significantly over time? I know you have worked in several different firms, but have there been some overarching changes in the way research has been done in your career?

Gray: Absolutely. Yes. It used to be that scientists did not expect to get rich. In fact, many people in science still don't expect to get rich. Most physicists and astronomers don't expect to get rich.

Computing, however, has turned out to be something that has huge value. People who are doing work in computing are creating things of huge value. It was a bit of a shock to me. When I left IBM it had a third of a million employees. It was a forty billion dollar a year company. Certainly it was able to pay for a huge amount of waste. The profit margins were not that extraordinary, so fundamentally the money was going into waste. I showed up at Tandem in 1980. Tandem was at that time about a 400-person company. You could see the whole thing directly. You could see how much wealth these very few people were creating. People would go to work and six months later they would ship a product and the revenue would start coming in and the numbers were just huge. That was the first time it dawned on me how much wealth was being created. In the last four or five years things got a little bit out of hand. I think people overvalued how much wealth was being created. But certainly it was substantial. Maybe it is not three trillion dollars, maybe it is only a half a trillion dollars—but it is still a lot of money.

Frana: Was there a proliferation of people who were interested more in the money than in the science?

Gray: No, but that is a good question. When I was working at IBM and even to this day, my motive is not economic. It is, 'This is interesting and if they pay me to do this that is great.' I think that motive is still widely held. There are lots of people with the intellectual curiosity motive and the craftsman motive of wanting to build things that are really neat. The Linux community is the classic example of this. But all of us our now are at least aware of the fact that intellectual property is in fact very valuable sometimes. You asked if the climate has changed? I think the pendulum is swinging back the other way, but recently the pendulum had swung very,

very strongly towards the goal: ‘How can this be used to make a lot of money?’ That makes for big changes in the research climate.

Frana: Research is getting more altruistic again?

Gray: The people who were fundamentally interested in ideas all along, continue to be interested in ideas. I think the people who were greedy have come to realize that there is actually a fair amount of work, and you actually do have to have an IQ. People look back on the previous generation and always romanticize the past. People look back on the golden age of Xerox PARC and Bell Labs. The simple fact is there are fewer industrial research labs by any metric per capita, per dollar of revenue, per this, per that. Oracle doesn’t have a research lab. EDS doesn’t have a research lab. To their credit, IBM still has a research lab—and I think they are using it better. There was a period when bean counters were in charge. They have turned it around. People who work there find it an exciting place to be, and don’t complain about the pressure. Microsoft has a new lab and it is a model. I think it compares favorably with Bell Labs at its best. Time will tell if Microsoft Research is a new Florence. I am not talking about in terms of its output or anything like that. The fact is that it is almost an exception. The other large software companies don’t have research labs. Computer Associates—they don’t have a research lab. You run down the list and it is a long time before you find one. Cisco had a research lab, but I don’t know what the situation is there at this point.

Frana: Was there a difference in the lab culture? You’ve worked lots of places. Were there different lab cultures?

Gray: Oh, very different. Yes. First you have to appreciate that AT&T was a very, very hierarchical place. Ken Thompson and Dennis Ritchie had freedom comparable to any group, but where I was working at Bell Labs I had almost no intellectual freedom. I was working on a project where they told me what I was going to do. Bell Labs was well organized—they had some places where there was lots of freedom and some places where there wasn't. IBM Research is similar. It has a whole spectrum from theoretical to applied technology. There is a spectrum of intellectual freedom, and it is pretty well managed. Microsoft is young enough that they don't have that structure. Microsoft tends to attract people who want to have impact. They have some people who are doing what I would call unfettered research. Most of the people actually are eager to have their ideas applied, and are out marketing them to the development divisions.

Frana: Is there risk-taking going on?

Gray: Well, the management certainly has been encouraging people to work on things that we don't know how to do. Speech is a longstanding research area for Microsoft Research. The community has made surprising progress. Language understanding is another long-term thing. Progress is actually tangible at this point, whereas for a long time it was just promising.

Frana: In the 1960s there was a lot of optimism and it faded. You and Gordon Bell are quoted at the end of the first chapter of John Seely Brown's *The Social Life of Information*.

Gray: Yes.

Frana: The quote is, I think it is from a 1997 report that you guys collaborated on. You say, ‘By 2047 all information about physical objects, humans, buildings, processes, will be online.’ The next sentence is, ‘This is both desirable and inevitable.’ That sounds like saying the best defense is a good offense. Are there people who are afraid of this?

Gray: I am terrified of it. Aren’t you? I mean, the fact that everything about you could be known, and that there are no laws to protect your privacy. The whole notion of privacy is not written into the Constitution. We don’t have a right to a private conversation. Period. This is a huge surprise to me. I mean they can get a court order and find out what you said, and what I said. You are recording this. The law about this recording is pretty vague. The whole notion of privacy is being eroded. People have credit cards, and now you know more or less where everyone is because you know what they bought. Cell phones are now going to have GPS. They are going to be able to track where you are to within a meter. Gradually, anonymity and privacy and the ability to have a private thought are disappearing. Imagine that you couldn’t have a private thought. Then imagine that somebody is watching and deciding which are the good ones and which are bad ones: ‘Jim is thinking a little strangely, I think we need to give him a little bit of lithium.’ Well, it is not so hard to imagine. People now have codes of behavior—one thing that allows us to have independence and so on is the fact that no one is looking most of the time. In the world that we just described, computers can in effect be looking all the time. Imagine that the Stazi had had good computers. Imagine if Hitler had really good computers. Fundamentally, it would have been very hard to hide a Jew if they’d had really good national identity cards.

Frana: The dark side of the personal Memex?

Gray: That's right. Inevitable? Nothing is inevitable. Bill Joy has been talking about taking a responsible view of technology. I think Bill is taking an alarmist approach. He believes you need to do something to wake people up and say, 'Hey, look where we are headed! Do you really want to be there when we get there? Is this the world you want to be in?'

Frana: Can the world be reduced to information? Can the mind be reduced to data processing?

Gray: We don't know. Hans Moravec believes it can. I believe it can. I don't see why not. Turing thought it could.

[Side 2:]

Frana: Just one final thought. To your mind is the world finite, or is it infinite?

Gray: Quantum theory would tell you that there are a lot of possibilities, but not an infinite number of possibilities. We are beginning to realize that our universe is finite, which is to say that the way that the expansion seems to be going, you go back fifteen billion years and you draw a sphere around us, there is a finite amount of matter in the universe that we can perceive.

Frana: There is an edge.

Gray: There is an edge, yes. Quantum space says that things can only be in certain states. The number is huge, but it in fact, is not infinite. It is such a big number that it is logically equivalent to being infinite. Where we are in the universe is lots and lots of zeros, hundreds of zeros behind it, but it is a finite number. So it seems. There was a piece in the *New York Times* recently about the future of the universe. It talks about how, I think a hundred trillion years from now, things are going to get very boring because everything will have gone inside of some black hole or other. All of our atoms will be inside some black hole or other and our universe will consist of what is inside of that black hole. Amazing huh?

Frana: Definitely. Should we get some lunch?

Gray: Yes.

Frana: Thanks, Jim.

Gray: You are welcome.

END OF INTERVIEW