An Interview with

WILLIAM McGEE

OH 331

Conducted by Philip L. Frana

On

12 November 2001

Los Altos, California

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

William McGee Interview

12 November 2001

Oral History 331

Abstract

William McGee is a retired senior programmer at the IBM Santa Teresa Laboratory. McGee received the AB degree in physics from the University of California at Berkeley in 1949, and the MA degree in physics from Columbia University in 1951. From 1951 to 1959, McGee managed the numerical analysis unit at the General Electric Hanford Atomic Products Operation in Richland, Washington. Between 1959 and 1964 he led systems programming and research at Ramo Wooldridge Corporation in Canoga Park, California. McGee joined IBM in 1964 as a staff member of the Palo Alto Scientific Center, where he specialized in physics applications, computer graphics, and database systems. In 1969 he received an Outstanding Contribution Award for work on Data Base/Data Communication requirements and strategy. In 1970 McGee joined the DB/DC development group in Palo Alto, where he managed the DB/DC Architecture Department. Other work in DB/DC development included performance evaluation, distributed data requirements and planning, and data dictionary planning and development. McGee retired from IBM in July 1992.

This is an oral history with William McGee conducted on November 12, 2001, in Los Altos, California. It is conducted by Philip Frana under auspices of the Software History Project, and funded by the National Science Foundation.

Frana: Thank you for participating.

McGee: You are welcome.

Frana: I was wondering if we could begin by having you tell me something about your early education and experiences and how you became interested in physics in the first place.

McGee: Well, I graduated from UC Berkeley in 1949 with a degree in physics. I went on to Columbia University and got a Masters degree in physics. Coming back to California I was looking for a job opportunity that would take me out of the reach of Selective Service. I did an interview with General Electric Company, which was hiring for the Hanford atomic works in Richland and I received a job offer from them, which provided me with that protection, so I took the job. When I was there I was under a rotational training program where they give you a choice of assignments and I selected physics as my first choice and I selected something called computing as my third choice.

Frana: Was it actually called computing?

McGee: It wasn't called computing. It was called numerical analysis or something like that. And I got my third choice. And so I went to work in their computing unit and I never left it. The work was so interesting and we were working with engineers and physicists and that sort of seemed to satisfy my need for the technical content of the job.

Frana: Were your parents scientists?

McGee: No. My father was a lawyer and my brother is a physician.

Frana: And your mother?

McGee: She was not working. She was a teacher many years ago.

Frana: And did you grow up in this area here?

McGee: Yes, I grew up in a little town up in the foothills called Sutter Creek. I went to high school there.

Frana: And how did you choose then to go to Berkeley?

McGee: Well, my whole family went to Berkeley. Both my father and my brother went to graduate school in the East, so that is basically why I went to graduate school in the East.

Frana: And you started out in physics then?

McGee: Right.

Frana: That would have been of interest to you in high school?

McGee: Yes, I was interested in engineering and physics.

Frana: You had a particularly good physics teacher or engineering teacher?

McGee: Yes. I had a very good physics teacher in high school. When I got into college I found that I was more interested in the math part of engineering than I was in the engineering part of it, so that is the way I drifted into physics I think.

Frana: And how did you end up all the way across the country at Columbia in ’51? That is quite a long way from home.

McGee: Well, like I said, both my father and my brother had gone there. We sort of had this family tradition of going East to go to graduate school. So that is why I was there.

Frana: Okay. Did you have a particular mindset that made you a good numerical analyst?

McGee: No, I don’t think so.

Frana: It is not a particular kind of mathematician that does very well in numerical analysis at this point?

McGee: I don’t think I am very strong at mathematics.

Frana: So then you went to work at the GE Hanford Atomic Product Operation, is that right?

McGee: Yes. Right.

Frana: And that was in?

McGee: 1951.

Frana: ‘51. Were you working on classified projects then immediately?

McGee: Yes.

Frana: What kinds of things were you involved in right away?

McGee: Well, there were many different kinds of applications that we were involved in. I was mostly involved in things having to do with reactor physics and reactor engineering, such as heat

transfer calculations, construction of tables for doing physics applications. This was before there was anything like personal computers to do your calculations with, so these engineers and physicists would bring their "problems," as they called them, into the computing unit and we would then program them for the machines we had at the time- which were the Card Program Calculator, and the 604 Calculator, and the 605 Calculator.

Frana: The IBM machines?

McGee: Right.

Frana: What I know about Hanford in the '50's is that corrosion on the fuel elements, the fretting was a particular problem – was that something that you would have worked on?

McGee: Right.

Frana: Reducing the corrosion?

McGee: Yes.

Frana: It is my understanding that eventually you got a 701 or 702 machine?

McGee: Right. The first large machine we got was a 702. And this was due mainly to the fact that most of the work that had been done in the computing unit was of a commercial type. We were doing the payroll, and the accounting, of various kinds for the plan. And the numerical analysis was just part of the work that we were doing. The 702 was eventually replaced by the 705, and then I think the next machine was a 7090. At that point, I sort of lost track of – that was in 1959, I think, in that period, and I lost track of it thereafter because I left Hanford in 1959.

Frana: Did it require lots of training to use these new [systems], the 702, and then the 7090? Did you have to take courses?

McGee: Yes. The courses were, in general, provided by IBM.

Frana: They would actually come in?

McGee: They would come in and teach us how to use them.

Frana: Were there IBMers on site then for a period of time?

McGee: Yes. Right. There were what they used to call 'customer engineers' who would be responsible for maintaining the equipment. They would also have system engineers who would help with the application work in developing general-purpose programs of one kind or another. So they were very closely involved with our operation.

Frana: I noticed in one of your papers that there is an example of the kinds of things that I suspect you used the 702 for at Hanford: payroll is one thing and radiation calculations?

McGee: Health sciences, yes. All employees were required to wear film badges and these were processed regularly and the data was collected and put through the computer.

Frana: But by and large most of it was keeping track of administration and payroll?

McGee: A lot of it was.

Frana: A lot of it was?

McGee: Yes.

Frana: Any idea? I mean, 70/30?

McGee: Probably that's about right.

Frana: Something like that? Scientific versus business application.

McGee: Right.

Frana: Now, you were the main contributor to the 702 File Maintenance and Report Generator System? Or was that a team effort?

McGee: That was a team effort.

Frana: And when does that come along? The mid '50's?

McGee: The mid '50's. My role in that was mainly as a person who recorded the results of the work.

Frana: How old were you at this time, in the mid-50's?

McGee: Twenty-five, thirty. And the reason that work became generally known was, I believe, an article, which I wrote for the *Journal of the ACM* in 1959. And the work itself was sort of an outgrowth of things, which had been going in other locations under the general rubric of formatted file systems. The military was heavily involved in that kind of a software system. So some of these ideas were just about ready to jell into the package.

Frana: What was the problem that you were trying to address in the File Maintenance Generator system?

McGee: Well, it was a programming problem. It seemed like every application required a large amount of programming effort to get it going, and it was our observation that many of these

applications had large parts of their processing that were generalizable. They all did sorting, and they all did file reporting, report generation and file maintenance, so those were generalizable operations.

Frana: So you could separate the files from the applications?

McGee: Well, it generally meant a partitioning of the application in ways, which you wouldn't normally think of doing, if you were programming from scratch. Or 'runs', as they were called then, were involved to get the whole job done. But the savings in programming effort was worth the additional machine time that the generalized approach required.

Frana: Now was this something that IBM was immediately interested in?

McGee: Very much, yes. And they were heavily involved in developing it too. The system was picked up by IBM and given to the SHARE organization, this is the IBM users group, and they maintained it for a while.

Frana: So is this the very first generalized report generator?

McGee: Well, it was one of the early ones. Like I say, there were other generalized programs that were going along at the same time.

Frana: So calling it the first is hard to say because it was just 'in the air', so to speak, but a very important one. Does this system then evolve into other systems down the line? I know quite a bit about the RPG Series from Minnesota, that is the Rochester contribution to the world.

McGee: Right. I think it is fair to say that the RPG was heavily influenced by this generalized system, but not totally. I think that work in that direction was going on independently.

Frana: Now Dan McCracken, was he working on this with you? Was he in your department?

McGee: He was in my department. But he was not involved in the development of the generalized system. He and I were there in the early '50's working on general purpose control panels for the CPC.

Frana: He has a nice photo of the two of you together in those days. I don't know if you've seen it?

McGee: Yes, I have. He left in about 1953 I think to go to GE at Evendale.

Frana: And what about Don Gowan, do you remember him at all?

McGee: Yes.

Frana: Did your time overlap?

McGee: Yes.

Frana: I know that he left. He was a mathematician and then left to go to the seminary I guess.

McGee: Right.

Frana: Do you remember him having this debate with himself about which way to go with his life?

McGee: No. No. I know he was very quiet; an introspective sort of person, but we never talked about theology.

Frana: Okay. Were there other Berkeley grads that had followed your same path? More or less that you knew from other places?

McGee: There were a lot of Berkeley students there but they were in other departments; engineering departments.

Frana: And I have to ask you this, did you ever meet General Groves while you were on duty at Hanford? I know that he didn't retire for – oh, maybe he had retired?

McGee: I think at that point he had.

Frana: Had he?

McGee: I think he had, yes.

Frana: I know he lived into then '60's. Did you actually work on site at Hanford?

McGee: Yes.

Frana: Some of the workers were exposed to radioactive iodine -131. Is that something you worry about today?

McGee: No. We were located in what was called the 700 area, which is in downtown Richland and the processing plants were located out by the Columbia River, some as far away as fifty miles. So part of my work in the later years took me out to one of these processing sites regularly, but at that time there was no major concern about contamination. And it is only in recent years that we have become aware of the real mess that we created.

Frana: Right. I think the early eighties was the first time there were some concerns raised about that. You don't have to answer this if you don't want to, but did some of your friends develop thyroid cancer?

McGee: Not that I know of. No.

Frana: That's good. Was Charlie Bachman around then?

McGee: Yes. Charlie at that time was working for General Electric in Phoenix, in the computer department. He was developing a system called Integrated Data Store. For some reason Charlie has always attributed the source of that system to the work that was done at Hanford, but I could never see the similarity between the two. IDS was, as you probably know, a system in which data was stored as linked records and the notion of linked records was something that came along after we had storage devices that would be capable of holding them. And we'd never had such storage devices at Hanford. All we had was magnetic tape and so the notion of a linked record would have been very strange at Hanford.

Frana: Because you had been thinking serially?

McGee: Serial. Serial processing, basically.

Frana: He remembers you fondly actually. He's quite a good friend of the Charles Babbage Institute where I work and he's been working on his memoirs and you are mentioned fondly in that work for your contribution to IDS, actually. He still firmly believes that your system was in part responsible for his success.

McGee: I don't know where he gets that idea.

[laughter]

Frana: Did you work then hand in hand with him on the specifications for 9-PAC?

McGee: 9-PAC was the SHARE version of the generalized system.

Frana: Okay. So you had many conversations with him then?

McGee: Well, some. I don't know what his involvement in 9-PAC was because that is too long ago to remember. Basically the way SHARE operated was that they would pick up packages from other companies, from companies that had IBM equipment, and they would make these available on a shared, no cost basis essentially.

Frana: There would be a routine library?

McGee: There would be a library. They would do a library function essentially.

Frana: And who would maintain that?

McGee: Some group in corporate headquarters.

Frana: So, did 9-PAC then differ much from the File Maintenance and Report?

McGee: No.

Frana: So let me see if I've got this straight. Did it not abandon the older punch card technology that stored the foreign key values in detailed records, rather than implying that their value comes from the master records that precedes them? Is that irrelevant to that leap?

McGee: The generalized routines and 9-PAC had the concept of a master record and detail records but that relationship was usually implemented by just sequentially storing them in the master detail order. Like I say, the liberation from that concept was something that was made possible by direct access storage devices.

Frana: And we are not quite there yet?

McGee: Yes.

Frana: So how then do you come up with the modern idea for schema in about '58, '59? This is about the right moment, right?

McGee: Well, actually that was sort of unrelated. In generalized routines we had the notion of a master file and we had the notion of making an explicit description of files, but that is what we call the descriptions. It wasn't until the late '60's, when I was working at IBM and I was trying to write an article about the history of generalized file processing. And there seemed to be a general acceptance of this idea of making separate descriptions, but there was no one term that we could use to refer to the abstraction of a description. So at that time I hit on the idea of using the word schema to represent this abstraction.

Frana: Does the definition of schema change over time at all? Has it remained basically what you just said or does it- sometimes these words tend to end up being co-opted and used for other purposes.

McGee: No, it has pretty much remained the same. There has been a tendency to use variations on that term, but I think the idea is pretty much the same.

Frana: And overall, just to sort of reiterate here, your idea is designed so that databases can ultimately be shared and be program independent? Would that be the ideal? If we were to abstract, and take it to its greatest extent, is that the ultimate objective?

McGee: That's really putting more meaning into that word than I think is really intended. I think the intent is merely to have a term that represents an abstraction of the structure or the format of a given collection of data. And every collection of data will have its own schema, in that sense. And that's basically all it is. The form that the schema takes is heavily dependent upon the overall model that is being used to represent the data, the main ones being the hierarchic, the

relational, and the network. Each of those models will have schemas, which will reflect that general style of the model.

Frana: Let's talk about your work at Ramo Wooldridge. I read this paper that you worked on the specifications and concept for the TRW 133? You weren't at TRW very long. Did that dominate most of your time?

McGee: I left Hanford in 1959, and went to work for Thompson Ramo Wooldridge in Canoga Park. And I was there for about five years, and then I came north to Palo Alto to work for IBM. I went to work there mainly to get experience in developing computer systems hardware. And of course I didn't have the engineering background to participate directly in circuitry design or anything like that. But I did get involved in the software for the systems and I did manage to pick up quite a bit of knowledge about the hardware technology of that era.

Frana: How did you come to leave Hanford then?

McGee: I just ran out of jobs; ran out of interesting things to work on.

Frana: It had been done; everything that you needed to have done was in place?

McGee: Well, the opportunity to learn more things wasn't there really and I was more interested in computing than they seemed to provide opportunities for.

Frana: Were you involved at all in the development of the Polymorphic Computer?

McGee: I was involved in writing proposals for the use of that system; that concept.

Frana: The RW400?

McGee: Right.

Frana: What were these systems? TRW was interested in supplying systems to NASA in the space race, and satellites and systems, was that what these systems were designed to do?

McGee: They were designed for command and control work. The TRW 133 was a commercialized version of the ANUYK-1, which was a hardened computer for use on board seagoing vessels. But their business was mostly in defense and I don't think NASA had been quite formed yet at that point.

Frana: Not quite. Air Force programs?

McGee: Air Force.

Frana: And then Glen Culler, was he at TRW at that point, or RW? I am not sure when the 'T' comes along?

McGee: RW was joined with the Thompson Products; I think it was in the early ’60’s.

Frana: And Simon Ramo had already left before you had arrived?

McGee: Yes, I believe that is correct.

Frana: Who were your colleagues there? People that we ought to know about?

McGee: Yes. My manager was Walt Bauer.

Frana: Oh, I know Walt.

McGee: He left and went on to form the Informatics Corporation.

Frana: Do you see him often?

McGee: No.

Frana: He’s got more energy than I do.

McGee: Yes.

[laughter]

McGee: He took with him a fellow named Werner Frank.

Frana: Oh yes. The fellow who did GIRLS- John Postley? Maybe he came from someplace else? Did you know John at all?

McGee: I haven’t heard that name in a long time.

Frana: Right. He lost his eyesight recently. I don’t know if this is the appropriate moment for this or not- but I asked Don Chamberlin this and we didn’t get a chance to talk about it in any detail. But he says that there is a direct connection between early database design and timeshare operations. Is it too early to be talking about that connection now in the mid-60’s? Were you working on timesharing systems when you were at RW?

McGee: Yes. Glen Culler I think was involved in development of online system usage. He was developing display technology that would permit shared use of computers. But the problems were not- the problems were more providing a robust kind of supervisory control of the systems and not so much the storage of data. At that time, I don’t think the data sharing problems had really surfaced, which is probably what Chamberlin is referring to.

Frana: And that would surface only as more and more users were linked?

McGee: Right. More and more users were linked into a series of systems rather than just using systems, for example, to develop programs.

Frana: Okay, so you were at RW for five years and then you left for IBM. And your decision to leave RW was based on?

McGee: The work was running out down there and I believe that division disappeared a few years later.

Frana: What was the division called?

McGee: RW Division of TRW.

Frana: Okay, I see. And you left about the same time that Walt did?

McGee: Yes.

Frana: And then you moved up to Palo Alto. And is that laboratory still in existence?

McGee: I moved up to an organization called the Scientific Center. IBM at that time had scientific centers in six or seven locations around the country.

Frana: At this time Santa Teresa was already around?

McGee: No.

Frana: This was before Santa Teresa?

McGee: This was in 1964. The Scientific Center in Palo Alto, which is where I was, was very closely linked to it the Stanford Computation Center and the Computation Center at the Stanford Linear Accelerator. And SLAC was one of the first System 360 installations on the coast and we were involved in setting that up. I have fond recollections of dashing in and out of the building with little decks of cards in our hands.

Frana: And that was the first task that was set upon you when you moved out to Palo Alto then, was carrying cards around?

McGee: Right.

McGee: No, I was involved in graphics application, very little having to do with database at that time.

Frana: You mentioned that in your resume here. What kinds of graphics applications did you get into? I was unaware that IBM was working on such things.

McGee: Displaying physics data like scatter diagrams of particle counter measurements.

Frana: So graphs and charts and tables and so forth?

McGee: Right.

Frana: Were you trying to display special mathematical characters in equations where you would have to use non-standard characters or anything like that?

McGee: No.

Frana: And who were the other people working? Do you remember the names of some of the other people that were doing such things? Anyone who has gone on to make that their career path?

McGee: Not really. I can't think of any. I still see people from the Scientific Center from time to time. They were mostly people with physics degrees who went to work for IBM. Eventually that Scientific Center concept proved to be too expensive to maintain, so that I believe it is no longer in existence.

Frana: So how did you recover your interest in database?

McGee: Well, it was through my association with a fellow named Bill Olle. Bill Olle at that time was working for RCA. Or was he working for Control Data, I think, in Palo Alto, and shortly thereafter for RCA in New Jersey. And he and I had a common interest in generalized file processing systems. At that time the CODASYL committee was organizing what they called a systems committee. And he was the chairman of that, and he invited me to join that committee.

Frana: And is that how you met Gio Wiederhold too?

McGee: Probably.

Frana: When did you first hear about the relational model then? Did you ever meet Ted Codd?

McGee: Oh, yes.

Frana: How did you hear about this?

McGee: Well, it was mainly through the papers that he published in the late '60's I guess. There was a group working at IBM, who - well, some in San Jose, and some in Poughkeepsie - that were becoming increasingly concerned with the programmers involvement in the physical details of the data storage. And so they came up with this notion of what they called an 'entity set model' which was very elegant. Of course it had not been proved by way of any implementation.

But I think Ted Codd saw the potential here for taking this concept and giving it a rigorous mathematical background basis. So that became the relational model.

Frana: Now did this require known technology or did the hardware follow on the heels of the idea?

McGee: The hardware solutions were not known. The hardware implementations were not known at the time it was conceived. It was more a matter of faith that hardware would be able to make up for what many people viewed as a very large potential for performance problems implicit in the model.

Frana: Right. That's important to me because I am really trying to figure out how much reluctance there was to make this change in mindset, if you will, and what the sources of that were. So there is the technological question, "Can it be done?" I've heard people complain that well Codd was highly mathematical. We read his paper and had no idea what he was talking about.

McGee: Well, that's part of it.

Frana: That's part of it as well? Was there inertia behind the network model?

McGee: Right. I recall one instance; down in, I think it was in San Diego, where I was going to a meeting of a local ACM group, and Richard Canning was there also. He published a journal called *EDP Analyzer* for a number of years. And his background was in business data processing. And Ted Codd was the invited speaker at this meeting and Ted got up to explain his relational model to basically a group of business processing people and on the way out they were all sort of shaking their heads. They didn't know what had been going on. I explained, I used an analogy, which might not have been very well appropriate. I said, "Well, think of a relationship. A relation is a punch card. It has fields that are laid out and it has a flat data structure. Each field on the card corresponds to an attribute in the relation. And Canning said, "Is that all it is?" He was very upset that such a simple idea was being couched in such elegant mathematical terms.

Frana: I've read the paper and there are some awful passages. There is a whole Greek alphabet in the original paper that doesn't seem to fit the other papers that are published around it. You can't learn the mathematical system of equations he is using from the other publications. But was Codd then his own chief popularizer, or is there someone else there that helped it along?

McGee: Well, he was the main driving force. He did get quite a bit of criticism from within the company, as well as outside the company. But I think the main value of what he did was to provide a theoretical framework in which a wide variety of data problems could be studied. This is not to say that it is the only framework that would have served the purpose, but it was the one that was used by many different people. And many, many papers had been published that studied such questions as data consistency, data protection, data security, all using his basic relational model as the theoretical framework. Many of the commercial database management systems which came into being in the '70's and '80's, either unconsciously or subconsciously or

consciously, adopted the relational model as their basic model. So you find things like Microsoft Access and …

Frana: Oracle.

McGee: Yes, the Oracle system, all have a very similar kind of data model, which is basically the relational model. So that's another valuable, fallout of that work. You can move easily from Oracle to Access if you just know the basics of what is going on in that model.

Frana: Now, I've never talked to Ted Codd, but I have talked to Chris Date. He is a freelance writer now I guess, and Date says, and I guess probably Codd agrees, that System R, which you did not work on directly yourself…

McGee: No.

Frana: …did not represent their ideas very well. Is it because they had a theoretical framework. When it was actually brought to life, it couldn't match the ideal?

McGee: Well, any time that you stop theorizing and start to develop a system you have to make certain decisions. I think that is probably what they objected to- the fact that research stopped with System R. System R was a running system, which eventually got commercialized as DB2. And DB2 of course went on to solve many of the performance problems which people thought were inherent in the model. But that is probably what Chris and Ted were concerned with, that they still saw many existing interesting problems, which could not be solved in that kind of a development environment.

Frana: Now it seems that the relational model gets used a lot in business applications, but is there any advantage to the relational model that would lead people to think that this is a great way of thinking about organizing scientific information? I mean, is there anything about the system itself that lends itself to business applications or is it just IBM is obviously a business data processing company more than a scientific applications company?

McGee: I don't think so. One of the big problems that people have expressed about the relational model is that the kinds of attributes that can be represented tend to be limited. But the commercial systems have overcome this by allowing attributes of any type or a large number of types like pictures and sound files.

Frana: Now let's get back more directly to your timeline. Most of your time in the mid-to-late '60's and early '70's was spent working on IMS/VS DB/DC for System 370? Is that right?

McGee: Right.

Frana: You spent a colossal amount of time working on that system?

McGee: I think the one application; the one package that I probably spent more time on, was the data dictionary system, which was an IBM/ IMS adjunct kind of thing. In fact it was an IMS application, but its main use was to enable installations to make dictionaries of their own applications data. And that system occupied quite a bit of my time in my later years with IBM.

Frana: And so, the IMS development, was that centered out here?

McGee: Yes, it started out in Los Angeles as part of IBM's relationship with North American Aviation. The so-called DL1 model of IMS was developed by a fellow by the name of Uri Berman, whom I still see occasionally. He is living in Palo Alto.

Frana: Now this is just a terminology thing for me, but is IMS/VS the same thing as the IMS Data Base and Data Communications as it is known now? Those are equivalent?

McGee: Right. Right. DB/DC is sort of a catchall phrase.

Frana: DB/DC, which you've written about.

McGee: Yes. IBM got interested in that in the late '60's when they realized that a lot of their branches were off developing generalized packages for use in database and data communication. And so they put together a team.

Side B:

McGee: IBM got concerned about various branches of each independently developing software support for database data communication applications. And one of these was IMS, and one was a package called CICS, and one was a package called GIS, which stood for Generalized Information System. And so they formed a task group to develop a strategy for the company and we met in Los Angeles for a year or so to develop a strategy in which those three were to be the initial IBM offerings in this kind of software support. And they became program products, among the first program products that IBM developed in the late '60's.

Frana: It was identified, as such, as a proprietary program.

McGee: Well, that would have been the first time that packages had been sold by hardware companies. Prior to that, they had just been given away.

Frana: So this is related to the unbundling decision?

McGee: Sort of.

Frana: Was unbundling the watershed it is now made out to be?

McGee: Unbundling took place really about ten years prior to this, so IBM was, at that time, under stricter operating control as far as the kinds of things that they could do.

Frana: Who are the clients then for IMS/VS? Were they government contracts mainly?

McGee: All sorts of customers: banks, insurance companies, manufacturing companies.

Frana: I read that the Apollo Project depended on IMS/VS to keep track of its data.

McGee: Well, that is probably true. I tend to think of that more as the origin of IMS. In other words, it was North American that was involved in the original development of IMS.

Frana: Now, but IMS isn't relational is it?

McGee: No.

Frana: It is hierarchical?

McGee: It uses the hierarchical model. It achieves the equivalent of the relational - well, not the equivalent - but the same kind of expressive power that you get with more complex organizations can be achieved through what they call the logical record concept. But it was not a really direct way of getting more complicated data structuring; it was more of an indirect way.

Frana: That would suggest that the whole world didn't have to go relational. Is it not so obvious? You know I am a young person, everything is relational so far as I know. In retrospect it seems like, how can you have these two incompatible models operating side by side? And the answer is?

McGee: Historical accident, I think.

Frana: Okay.

McGee: The fact that you have so many creative people involved in thinking about these problems that you are bound to come up with more than one way of doing something.

Frana: Let me ask you some fairly, I think these are very simple questions, but they are things that I have yet to have recorded on tape adequately. How do you prevent two users from simultaneously accessing the same data segment? Would that be the right way of saying it, in a database? How is that achieved?

McGee: It is basically done through locks, when one user wants to access a segment or a record he will make a request for a lock and if the lock is granted then no one else can access it at the same time. And of course the whole technology is developed around this problem, what is the best way to achieve locking?

Frana: Because two people can ask for a lock at the same moment I suppose too?

McGee: Right. And they can get into a deadlock situation where A wants B, and B wants A and they are both waiting for the other to release it. The English call this the “deadly embrace problem.” But the people at IBM Research really did the pioneering work in this area. The paper which Chamberlin contributed on the use of predicate locks is probably the most widely cited paper in that area.

Frana: And data recovery, restart recovery procedures? How are they achieved and are there various different methods for achieving that?

McGee: There are different methods for doing that also. Most of them involve keeping a journal of activity that can be used if necessary to recover an earlier state essentially by backing out the changes that you have made to an earlier checkpoint.

Frana: Okay. So in essence you would go back to a moment where you are sure everything has been locked?

McGee: Your system will take system checkpoints periodically and if the system crashes following a checkpoint, then you use the journal to backup to the most recent checkpoint.

Frana: And the System 370, did it also have very sophisticated data security protocol?

McGee: You mean basic to the hardware?

Frana: Yes. Or is there no such thing.

McGee: Well there are. I am sure there are security features in the hardware having to do with memory protection. But the larger scale protection and integrity has to come through the software.

Frana: Thanks for answering those questions, because you seem to have a very special gift for teaching other people that I have seen in the numerous pieces that you have written. I noticed that you have a special ability to create these kind of synthetic pieces that achieve clarity. How did you acquire such skill?

McGee: I’ve always been interested in writing and doing what I call expository writing. I don’t know where I got the interest.

Frana: Okay. Several of the pieces that you wrote, especially in the seventies, seem to have been designed to teach other people. Was that your charge?

McGee: No. I am glad that it had that outcome but that really wasn’t the main impetus. I think one thing that may have inspired me was the educational material, which we are all exposed to. When we are in the process of learning we don’t really recognize the inadequacy of that material. I look back on some of the textbooks that we used in college, I picked them up years later and I realized that they are really very poorly written. At the time of course, they seemed like the

gospel, you know. And the same was true in the software industry. A lot of the manuals that were produced were very badly written. And I guess that is still true today.

Frana: Perhaps, for different reasons now.

McGee: Probably.

Frana: We have technical writers who translate.

McGee: Well, that's part of the problem. But it may have been just a desire to overcome this kind of bad writing that was going on, bad documentation.

Frana: I remember reading about five or six years ago in *Fortune* magazine that IMS was still used in ninety percent of all Fortune 1000 companies. Is that still true?

McGee: I don't really know. I have been gone from IBM for ten years now. I have really not followed the course of the products. I know that DB2 is still very much a contender, because of what I see in *The Wall Street Journal*. But I never see IMS advertised in *The Wall Street Journal*.

Frana: But your guess is that it is probably still used widely in legacy systems.

McGee: Probably. Yes.

Frana: Nothing else.

McGee: Right.

Frana: What's the relationship between IMS and MVS then?

McGee: Well MVS is an operating system. It is like Windows. And IMS is an application that runs under the operating system, much like Access would run under Windows.

Frana: So there is really no relationship other than that is an operating system that supports IMS.

McGee: Right.

Frana: So then in the '80's you were working on…

McGee: IMS, the DBDC Dictionary. I worked on DB2 for a while helping develop test cases.

Frana: And you moved to Santa Teresa in what year?

McGee: Well, the organization moved in 1974. I moved out of the Science Center into what was then known as the System Development Division of IBM, which was developing these large software and hardware systems. Enterprise solutions I guess you would call them today. And

then a few years later we physically moved from Palo Alto down to Santa Teresa to a brand new facility.

Frana: And were you commuting all the way down to Santa Teresa from here?

McGee: Right.

Frana: It is maybe not as far as it seems.

McGee: It was far, 35 miles.

Frana: Some of the people who use these transcripts are going to want to know something about your leisure activities, what makes you, you. You play the piano I know; you sing in a quartet. Do you have other leisure activities that consume a lot of your time?

McGee: Well, one of the things I did after I retired was to go back to school. I went back to San Jose State to study music and got a bachelors degree with a specialty in composition. And so I've done some composing since then, arranging works for the quartet, the choir.

Frana: A remarkable number of you computer scientists are musicians. That can't be a coincidence. Is there something very mathematical about composing?

McGee: Not really. A lot of people have tried to draw some parallels, but any mathematics which tends to get used results in a very stilted kind of music which people don't really appreciate.

Frana: Do you worry about the present or future of computer science?

McGee: No. Computers seem to be a very ingrained part of our civilized world now. It is not something that is going to go away certainly. It could lead to some problems in terms of security and individual privacy but I don't see any insurmountable problems.

Frana: Let me ask you a final question here. Did you ever reach a crossroads where you could've explored something else; that you could've done other things with your life? Were there missed opportunities?

McGee: Not so much missed opportunities. About ten years before retirement I began to be very much less enthusiastic about the work that I was doing and if it had not been so late in my career I think I would have switched to something else. There were some assignments that I had that I wasn't really well suited for, which I stuck out just for the good of the team.

Frana: Was it the problem that you were forced into upper management against your wishes?

McGee: No, I really had to restrict that kind of a tendency. I was a manager for a while until I realized that was not my forte. So I had to make explicit requests to be left in technical positions.

Frana: Did IBM culture change in the ’80’s, where your gifts weren’t as well utilized?

McGee: I don’t think so. I think the best part of my time with IBM was when I was involved with the Systems Committee working on the DB/DC strategy because the work was new then and not really well understood and we tried to put some form onto the work, and the systems committee, as you probably know, published a couple of books which have been widely used to understand the early systems, how they evolved. And I think that is probably my most important contribution while I was with IBM.

Frana: Well, thanks so much Bill.

McGee: Okay, Phil.