Minutes of SCC meeting of April 26, 2006

by Nadia Ilyin

Attending

Al Kossow	Henry Gladney	Paula Jabloner
Bernard Peuto, Chair	Kathe Gust	Paul McJones
Bill Selmeier	Linda Johnson	Paula Newman
Dick Guertin	Michael L. Powell	Phil Gust
Gardner Hendrie	Nadia Ilyin	Ronald Mak

Presentations and Discussions

Al Kossow, Software Curator

Al Kossow was introduced as the Robert M. Miner Software Curator on the museum staff. Al gave a presentation on the present state and future directions of the collection. Much of the collection is post-1995, and there is not much software from the original (Boston) museum. The software has been whatever people gave, with no particular direction. The bulk is consumer software, PCs, floppies, and packaged applications.

More interesting and rare is the non-consumer software. There is little before 1975 in machine-readable form. The items of interest are noted in the handout. Some of this has been read by the restoration team, as noted in the handout.

Directions for the future include switching from reactive mode to getting things from people. Items from the mid 1970s and earlier are very hard to find. Any SCC contacts with people and old user groups should be used. The focus should be on:

- Software for machines that we have and are going to restore (diagnostics, operating systems, applications).
- Software for historically significant machines that we don't have (similar items)
- Copies of what we know about that is on the Web, as a backup in case those sites disappear when people die or become unable to maintain them any longer.

Al gave several examples of what we have along these lines in his handout.

There are also "latent" collections that are known to exist, but the collectors are "sitting on them" because they want them to restore their own machines. In this category, the MIT/AI Lab tapes are being read onto disks. The plan is to go on with what we are doing, but to be more selective. Not much budget is needed for the recovery lab. We do need the ability to read seven-track magnetic tapes. We have the only two contacts (outside the museum) who can do this, and we are satisfied with their work. Bernard Peuto commented that the museum should also have the means to read these tapes, nevertheless. Bernard and Ronald Mak commented that once the tapes are read, one must interpret them, and the ability to do so will go away soon.

Al pointed out that these magnetic tape media are fragile, also.

There was a question about Olin Sibert's work on the Multics collection project. Bernard commented that Olin is doing all the work and will give the collection to the museum, but we cannot expect other collectors to be as generous. CDC is difficult because of license issues; but Al is not too worried about this. He is more concerned about EMC. There is someone who has his own collection of Data General stuff, but there are "license hassles" potentially due to the axcquisition of DG by EMC.

For software archiving, the plan is to get a machine with a RAID. There is a local copy of bitsavers.org.

The group discussed sources that they knew who could read old video magnetic media. No particular names were given.

Paula Newman: ACM Book Initiatives

Paula Newman gave a presentation on her comparison of the CHM's current book collection with the ACM list of 600 favorite books, and her resulting evolving perspective on the process and standards for the book collection. She distributed her slides to the mailing list after the meeting; they are not repeated here.

Some interesting questions she considered and points she made were

- What does it mean to "have" a book, with respect to editions, versions, etc.?
- Why should we collect the books? It is more than just the history of individual systems, but the history of underlying ideas. This is harder to do with software than with hardware.
- There were influential ideas in software that were never coded, or were only coded long after their origin.
- The publications are jumping-off points: they occur before or between software implementations, or they taught how to write software.
- Initially we want to collect the older books, some rare and some not. (Paula gave several examples.)
- We also want biographies of the authors, and why the books were written. These are best obtained from the authors themselves, or from their colleagues, and some of these people are dying.

Paula suggested that collecting involves:

- selecting significant publications; focusing on the older ones
- obtaining context entries for each (author bio, summary of content, significance)
- providing online versions if copyright permits
- perhaps interesting academics and their graduate students in "adopting" a book and its description as a project.

The group discussed the problem that many existing libraries are "downsizing" their collections. We do not want to duplicate what others do; but we do want to get the materials that may be lost in this process.

Some computer science faculty no longer know or teach the history of computing, so the academic adoption approach may not work.

The group agreed that we need a shorter list of the books we want. There was some discussion of various ideas to achieve this. There was consensus that we need to set up guidelines on how to collect the books and the correct information about the books. Paula suggested that we choose five books and do a pilot project.

Paul McJones and Phil Gust: What Does "Done" Mean?

Bernard had asked Paul and Phil to put together their ideas on determining when we have finished a collection effort.

Paul spoke first and described his pilot project of collecting FORTRAN and Lisp artifacts. He distributed his slides to the mailing list after the meeting.

Paul's original goal was "to see what was possible": to pick a project and "see how hard it is." He decided to go after the original FORTRAN compiler. He "cast a lot of nets, and the fish didn't bite fast." After a year, he added the original Lisp system ("the un-FORTRAN") as a second, contrasting test case.

After two years on FORTRAN and one on Lisp, his remaining goals are

- Getting the collection into the museum archives
- Expanding the IBM search to related systems (704 through 7094)
- And "a bunch of other software" (listed on his slides)
- Land the Stoyan collection of Lisp documents

Paul described how he went about his search, mainly through an expanding tree of personal contacts and networking, and Web searching.

How do you declare success or failure? "It never stops." But the rate of change slows.

What happens when we are done? Make sure it is archived; communicate process and results to others.

The group asked Paul what worked and what didn't work. The most successful tool was his blog about his collecting progress. People responded by writing to him about what they had.

There was a question about interviewing for oral history. The museum is collecting oral histories and has the right equipment, but wants SCC to supply additional interviewers.

Phil Gust then spoke about the 18-month NLS/Augment project. He had distributed his slides to the mailing list before the meeting. Kathe Gust also worked on this project. The original goals are detailed in his slides, but included capturing and preserving the last running system. (He noted that Doug Englebart is 80 and not well.) Another goal was tracing the legal chain of ownership, which was difficult. An overall goal involves providing a basis for scholarship about the system and about more general questions such as

- What did a mature system programming language look like in the 1960s?
- How did its writers do that when 32K was a lot of memory?

Accomplishments in the project have been extensive (see slides). Among remaining goals are the collection of further materials, their preservation at the museum, and the organization of an online exhibit at the museum.

Phil learned that "lawyers are absolutely the wrong people to talk to" about getting rights to software that a company may or may not own after a chain of acquisitions. One should talk only to a very senior person in the company, other than the CEO, concurred Bernard.

It will be important to make recommendations and create white papers on how to do these processes. The process [of collecting] is as important as the outcome. Such papers could be published externally.

There is a need for orphan-work legislation changes; the museum or SCC could lobby for those.

To answer Bernard's original question as to when we are done, Phil said that we are done when we have the core, "one or two sigma on the bell curve," and when others can pick up where we left off.

Phil also announced that BayCon 2006 (S.F. Regional Science Fiction and Fantasy Convention, May 26 through 29) needs volunteers for two panels suggested by Phil and accepted:

Collection Historical Software
Orphan Works

Contact Phil at gust@acm.org to participate.

Bernard Peuto: Directory of Software Web Sites

Bernard Peuto presented a new project of the SCC: creating a directory of related software Web site. The intent is to have a set of pages on the SCC Web site that will document these links, not merely list them. We shall start with the links that are already present in existing project information such as that done by Phil Gust and Paul McJones. This list of links will be annotated and have search capability. An initial version will be done to demonstrate style and content and it will be reviewed for new suggestions. It should be a collaborative effort by SCC members.

Long-term goals are

- To save the information
- Possibly to mirror or host other sites that might be endangered by the age, ability, or interest of their authors
- To be the "Michelin Guide" of software collection and history Web sites, perhaps even with a star ranking system. (The group preferred to have readers do the ranking, rather than SCC.)

As Paul said, it is less important at this point to worry about taxonomy or metadata than to get these before they are

gone.

Bernard Peuto: Recruiting

Bernard reported that he has begun a focused, personal effort to recruit volunteers to carry out the actual tasks that the SCC members have wanted done. It is not enough to sit back and hope that people will find us. He mentioned several volunteers whom he had recruited so far.

The meeting adjourned at 7:30.

Upcoming Meetings

Day	Date	Time	Conf Room
Wednesday	May 17	5:30 - 7:30 pm,	Hopper
No June Meeting*			
Wednesday	July 19	5:30 - 7:30 pm,	Hopper
Wednesday	August 16	5:30 - 7:30 pm,	Hopper
Wednesday	September 20*	5:30 - 7:30 pm,	Hopper
Wednesday	October 18	5:30 - 7:30 pm,	Hopper
Tuesday	November 14*	5:30 - 7:30 pm,	Hopper
No December Meeting.			

* Note changes to previous schedule