

SAILDART Prolegomenon 2016

Author: Bruce Guenther Baumgart

January 2017

“History will be kind to me, for I intend to write it myself.”

– Winston Churchill

Abstract

A prolegomenon introduces a work of considerable length, which in this case is the digital content of the DART backup tapes from the first Stanford Artificial Intelligence Laboratory, SAIL. DART, Dump And Restore Technique, tapes were recorded between 1972 and 1990. This paper provides links into SAILDART, its provenance and an exegesis. This paper includes guidance for running PDP-10 software and for displaying 3-D models of the building and grounds. This paper concludes with notes on digital preservation. The style of this document varies from technical manual to hippie hacker memoir while its media formats include white paper, web pages, pamphlets and a picture book for Les Earnest's coffee table, which once was a hard disk at SAIL.



Copyright ©2016 All rights reserved except those granted by the licence below.

SAILDART Prolegomenon 2016 by Bruce Guenther Baumgart is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. Based on a work at www.saildart.org/book/2016.

Forward

Someone, other than the author, tells readers why they should read this book.

Preface

How this book came about.

This book is a time capsule marker that points at the content of the backup tapes of the lab where I did my graduate school work. Looking at the SAIL-DART archive as a tiny SD memory card or as the top page at www.saildart.org does not have the same impact as a white paper or as a glossy coffee table book. How I became a digital archivist. How my PhD thesis adviser, John McCarthy, asked me to take up the DART tapes, will be answered.

Introduction

This whole book is introduction to SAILDART.

The wide subject is Computer Software History. The narrow subject concerns what can be found on the backup tapes of the first Stanford Artificial Intelligence Laboratory. Unlike Julius Caesar, I shall write this history in the first person as I saw it as an A.I. graduate student, and later as I preserved it as a storage systems research associate at IBM, as a system administrator at Internet Archive, as an operations manager at the search engine Cuil, and now as a retired software engineer.

Contents

Contents	4
I Context of SAIL and DART	11
<i>The first four chapters set SAIL in its time and place introducing the people who worked on projects under the banner of Artificial Intelligence at Stanford. The fifth and sixth chapters depict the hardware and software that wrote the SAIL files that are on the DART tapes.</i>	
1 Time – Now 2016 and Then 1974	15
1.1 1963 start of the Stanford A. I. Project	16
1.2 SAIL Autobiography	18
1.3 1972 Rolling Stone and DART reel #1	27
1.4 1974 time portal	27
1.5 1979 End of funding and the move to Campus.	29
1.6 1990 Final DART reel is written.	29
1.7 1991 SAIL.Stanford.EDU is decommissioned.	29
1.8 Recap SAIL Time Period	29
1.9 About Time	31
1.10 Exercises	31
2 The Place – You are Here	33
2.1 Donald C. Power Lab	33
2.2 Margaret Jacks Hall	39
2.3 Wm. Gates Building	39
2.4 Polya Hall	39
2.5 Modeling for mobile robots.	40
2.6 Stanford University	41
2.7 About Place	42
2.8 Exercises	42
3 People and PRG codes	45
3.1 Significant SAIL people circa 1974.	46
3.2 PRG code related tables	51
3.3 Linking PRG codes to USER names and dates.	52

CONTENTS	5
3.4 People at the November 2009 SAIL reunion	53
3.5 People at the 25 March 2012 JMC Celebration.	55
3.6 Dead People, Gone and HEAVEN.	55
3.7 The next big reunion 2017	55
3.8 more digital images	56
3.9 SAIL Book of the Dead.	56
3.10 Exercises	56
4 Topics, Projects and PRJ codes.	59
4.1 SAIL project code preview	59
4.2 SAIL Topic Overview	60
4.3 SAIL file system project codes	60
4.4 SAIL accounting system project codes	60
4.5 Project areas with archive specialists.	60
4.5.1 PUB	61
4.5.2 TEX	61
4.5.3 LISP	61
4.5.4 Music	61
4.5.5 PARRY the Paranoid Simulation	61
4.5.6 Spacewar	61
4.5.7 Mathematical Theory of Computation	61
4.5.8 System and WAITS	61
4.5.9 SUDS Stanford University Drawing System	61
4.5.10 Foonly	61
4.5.11 36-bit computers	62
4.5.12 The Stanford CART	62
4.5.13 Earnest : FINGER and FONDLE.	63
4.5.14 Frost : AP news.	63
4.5.15 Gorin : DART, DSKUSE, ACCT, RALPH, LOGIN.	63
4.5.16 Baumgart : GEOMED, CRE, EDFONT, XIP, XAP.	63
4.6 Stanford University Course Work	63
4.7 Orphan Project Areas	63
4.8 Social Studies	63
4.9 Gender, Race and Class Topics	63
4.10 Exercises	63
5 Hardware	65
<i>This chapter is a triplet: outer system diagrams span two decades, middle surveys the major devices in 1974, finally inner is wizard craft for the chapter-17 re-enactment.</i>	
5.1 Block Diagrams	66
5.2 1974 Devices	69
5.2.1 CPU	69
5.2.2 CTY	69
5.2.3 SYS	70
5.2.4 DKB	71

5.2.5	III	71
5.2.6	DD	72
5.2.7	LPT	72
5.2.8	XGP	73
5.2.9	HAND EYE	74
5.2.10	CART	74
5.2.11	KLUDGE BAY	74
5.2.12	LIBRASCOPE	74
5.2.13	MAG TAPE drive devices MTA and MTU	74
5.3	Wizard Hardware Lore	75
5.4	Exercises	77
6	Software	79
6.1	Backup Utility: DART	80
6.2	Operating System: SYSTEM and WAITS	81
6.3	File system repair: RALPH	81
6.4	Help	82
6.5	Login and Logout	82
6.6	Accounting ACCT *SPY*	82
6.7	Assembly language: MACRO, FAIL and MIDAS	82
6.8	The LOADER	82
6.9	The Text Editors 'E'	82
6.10	RPG - Rapid Program Generation	83
6.11	Software content payload preview	83
II	Content of SAILDART¹	85
<i>The next four chapters, 7 to 10, are archive technical. There is one visual moment that justifies two photographs to document moving the magnetic tapes from Gates to Green. Chapters 11 and 12 will be heavily illustrated for a coffee table edition, arrays of thumbnails will suffice for this white paper edition.</i>		
7	Access, Privacy & Search	89
7.1	Public Access	90
7.1.1	Access by Canonical URL	90
7.1.2	Access from programmer Home Page	90
7.1.3	Access by Date	90
7.1.4	Access by Serial Number of content blob	91
7.1.5	Access by pathname	91
7.2	Copyright and Ownership	91
7.3	Privacy, Courtesy and Ownership	91
7.3.1	Who guards the guardians?	91
7.3.2	Stanford Research Policy Handbook	91

¹The banner means “A Discussion about the Content of the SAILDART” and where you can find the actual content.

7.4	Search	92
7.4.1	External Search Engines	92
7.4.2	Internal search mechanisms	92
7.5	Digital Curators	93
7.6	Exercises	93
8	Provenance	95
8.1	Custody	96
8.2	Fidelity	98
8.3	Prejudice	99
8.4	Infidelity	99
8.5	Exercises	99
9	Data Remix - Exegesis	101
9.1	Narrative Description of the DART format.	102
9.2	DART formats Illustrated.	106
9.3	Recipe for cooking DART data into GNU/Linux files.	109
9.4	Program text of /home/sail/undart-2014.c	111
9.5	Program text of DART.FAI[TAP,REG]	111
9.6	Program text of DSKUSE.FAI[ACT,REG]	111
9.7	Remove DART packaging from SAIL file data	111
9.8	SAIL text into UNICODE UTF-8.	113
9.9	Disassembling DMP executable binary into text.	113
9.10	SAIL software highlighting into cross-referenced HTML.	113
9.11	Exercises	113
10	Meta Data - Taxonomy	115
10.1	SAIL file system metadata	116
10.2	DART tape record metadata	116
10.3	Text verses Binary	116
10.4	Role of the Digital Curator	116
10.5	SAILDART cooked metadata	116
10.6	Exercises	117
11	Corpora Data Sets - dot • codes	119
11.1	• DMP	120
11.2	• FAI and • {nothing} and • MAC	120
11.3	• LSP and • LAP	120
11.4	• PUB and • TEX and • DOC	120
11.5	• MSG	121
11.6	• SAI	121
11.7	• DRW	121
11.8	• XGP	121
11.9	• PLT	121
11.10	• DAT	121
11.11	• FNT fonts	122

11.12 Generic • TXT • LOG • LST • OLD • TMP	123
11.13 • DMD	123
11.14 • REL	123
11.15 Languages • F4 • FOR • PAS • ADA • C • H	123
11.16 GEOMED files • B3D • CAM • CRE	123
11.17 SUDS Computer Aided Design Corpus	123
11.18 Digital Images Corpus	123
11.19 Audio Sound Corpus	124
11.20 Extension EXT code Theory	124
11.21 Large, Medium and Small collections	124
11.22 Document Sampler	124
11.23 Exercises	125
12 Non-DART Collections	127
12.1 Scanned papers and books	128
12.2 Scanned photographs	128
12.3 Digitized SAIL film	128
12.4 Post DART video	128
12.5 Inventory of SAIL documents	128
12.6 Inventory of SAIL photographs	128
12.7 Inventory of physical SAIL objects	128
12.8 Exercises	128
III Containers for SAILDART	131
<i>Cicero used alliteration to push power points. Preservation, Publication and Participation are in chapters 13, 14 and 15; the World Wide Web is in chapters 16, 17 and 18.</i>	
13 Preservation	135
13.1 Disk storage at SAIL	136
13.2 Preserving SAIL files	136
13.3 Preservation Strategies	136
13.4 Passing the Baton	137
13.5 Exercises	137
14 Publication	139
15 Participation	141
15.1 Form of Organization: Karass.	142
15.2 Participation Role: Digital Curator.	142
15.3 SAILDART Institute Job Descriptions	143
15.4 People I wish to contact	144
15.5 Exercises - this time for the Reader.	145

CONTENTS	9
16 Static WWW	147
17 SAIL System ReEnactment	149
<i>This chapter reports the work in progress on my quest to run all the old SAIL software via the SAILDART.org web site. Like a 19th century railroad steam engine, or an 18th century tall masted sailing ship, running the 1974 SAIL mid 20th century time sharing system is a major stunt that I have not adequately achieved yet.</i>	
17.1 SAILON-76	152
17.2 Keyboard	152
17.3 Exercises	153
18 Road-map Atlas	155
18.1 My Goal Immediate	156
18.2 Tools and Skills	156
18.3 Patrols and Sorties	156
18.4 Maps	156
18.5 Priority: Decision verses Recreation	157
18.6 Exercises	157
Bibliography	161
Bibliography	163
Index	165

Part I

Context of SAIL and DART

Part I Context of SAIL and DART**PART**

	chapter 1	
2	when	
3	where	
4	who	
5	what	
6	hardware	
	software	

*The first four chapters set SAIL in its **time** and **place** introducing the **people** who worked on **projects** under the banner of Artificial Intelligence at Stanford. The fifth and sixth chapters depict the **hardware** and **software** that wrote the SAIL files that are on the DART tapes.*

When, where and who did what in building computer hardware and writing computer software at the first Stanford Artificial Intelligence Laboratory.

Chapter 1

Time – Now 2016 and Then 1974

1

SAIL was the lab. DART was the backup program which wrote a final set of 229 reels of *permanent* tape. SAILDART is a digital archive propagating copies of the SAIL content since 1998. The time span recorded on the DART backup tapes is 1972 to 1990. The year 1974 is the earliest with sufficient computer records to *now* run the software as it was *then*.

1.1 1963 start of the Stanford A. I. Project

The ARPA funding for the Stanford Artificial Intelligence Project started 15 June 1963 with six people¹ in Polya Hall on the north edge of the Stanford Campus. The Project was equipped with an 18-bit PDP-1 computer. In 1966, the A. I. Project moved into the D. C. Power building at 1600 Arastradero Road. Perhaps the building itself caused the name to morph from “Project” to “Laboratory”. With further ARPA support, SAIL upgraded to a 36-bit PDP-6 computer. Computers now are commonly 64-bits wide.



Dave Poole working with the PDP-10 at the Stanford A. I. Lab.



Poole and Dave working with the HAL-9000 in a movie.

Figure 1.1: Artificial Intelligence in the late 1960s.

¹The six might have been McCarthy, Feldman, Feigenbaum, Russell, Peiper and Poole.

Time line 1st Stanford A. I. Lab

- **First decade+.1, 1963–1974**

1963-06-15 SAIL founded by Professor John McCarthy.
 1966 SAIL moved off campus to 1600 Arastradero Road.
 1966-06-06 PDP-6 computer installed.
 1968-??-?? PDP-10 KA computer installed.
 1972-10-18 Spacewar competition reported in Rolling Stone 1972-12-07.
 1972-11-05 First DART tape#1 written.
 1974-07-26 For this date, a software re-enactment is available

- **Midway half decade, 1975–1979, first A. I. funding winter**

1976-07-05 PDP-10 KL computer goes online (arrived on 1976-3-31).
 1979-11-01 SAIL moved to Margaret Jacks Hall
 1980 SAIL is merged into the Computer Science Department and CCRMA inherits the PDP-6 and the D. C. Power building.

- **Final decade+.2, 1980–1991**

1986 the D. C. Power building is demolished.
 1990-08-17 Final DART tape#2984 low and tape#3228 high density.
 1991-06-07 Final E-mail message from the SAIL Timesharing System.

The 2nd Stanford A. I. Lab opened in 2004 under director Sebastian Thurn. For this paper, the bare acronym, SAIL, refers to the 1st SAIL, except when context distinguishes SAIL the Laboratory, from SAIL the Programming Language, from SAIL a Stanford.EDU hostname. The SAIL host server outlived its laboratory by a decade hiding in a basement corner of the Stanford Computer Science Department building. The SAIL-WAITS Time Sharing started several years before and lasted a year after its 18 year DART tape recording window.

Numerologically auspicious the SAIL PDP-6 computer passed its acceptance test on the date 6/6/66. According to Les Earnest, the 36-bit time sharing era at SAIL endured for exactly 25 years, from Monday 6 June 1966 to Friday 7 June 1991 when at a couple of minutes before 9 pm PDT many of us received a first person, final biographical note, a last testament, from the A. I. person who signed off as “From: SAIL Timesharing System”.

Please read the following long quotation adapted from that final E-mail message which foreshadows what can be found in the DART backup records. *Or you may fast forward to figure 1.5 on page 23 to study early digital pictures*

of a nude female.

1.2 SAIL Autobiography

Date: 07 Jun 91 2056 PDT
 From: SAIL Timesharing System <SAI@SAIL.Stanford.EDU>
 Subject: life as a computer for a quarter of a century
 To: "@BYEBYE.[1,SAI]">@SAIL.Stanford.EDU

TAKE ME, I'M YOURS

The autobiography of SAIL²

I've had a very full and adventurous life. At various times I have been the world's leading research computer in artificial intelligence, speech recognition, robotics, computer music composition and synthesis, analysis of algorithms, text formatting and printing, and even computer-mediated psychiatric interviewing. I did have some help from various assistants in doing these things, but I was the key player.

I developed a number of new products and founded a string of successful companies based on the new technology, including Vicarm, Foonly, Imagen, Xidex, Valid Logic, Sun Microsystems, and Cisco Systems. I also gave a major boost to some established firms such as Digital Equipment and Lucas film. What did I get from all this? No stock options. Not even a pension, though Stanford is still paying my sizable electrical bills.

I was always good at games. For example, I created the advanced versions of Spacewar, which spawned the video games industry, as well as the game of Adventure and I was the computer world champion in both Checkers and Go.

I invented and gave away many other things, including the first spelling checker, the SOS text editor, the SAIL compiler, the FINGER program, and the first computer-controlled vending machine. Note that my name has been taken by the SAIL language, the SAIL compiler, and the laboratory in which I used to live. Just remember that I was the original Stanford Artificial Intelligence Laboratory.

Beginnings

I was born on June 6, 1966 at the D.C. Power Laboratory Building in the foothills above Stanford. I remember it well—the setting was beautiful, in the middle of horse pastures with views of Mt. Tamalpais, Mt. Diablo, Mt. Hamilton, Mt. Umunhum, San Francisco and the Bay, but the building itself resembled a flying saucer

²The claim that there was a ghostwriter is reviewed after the autobiography.

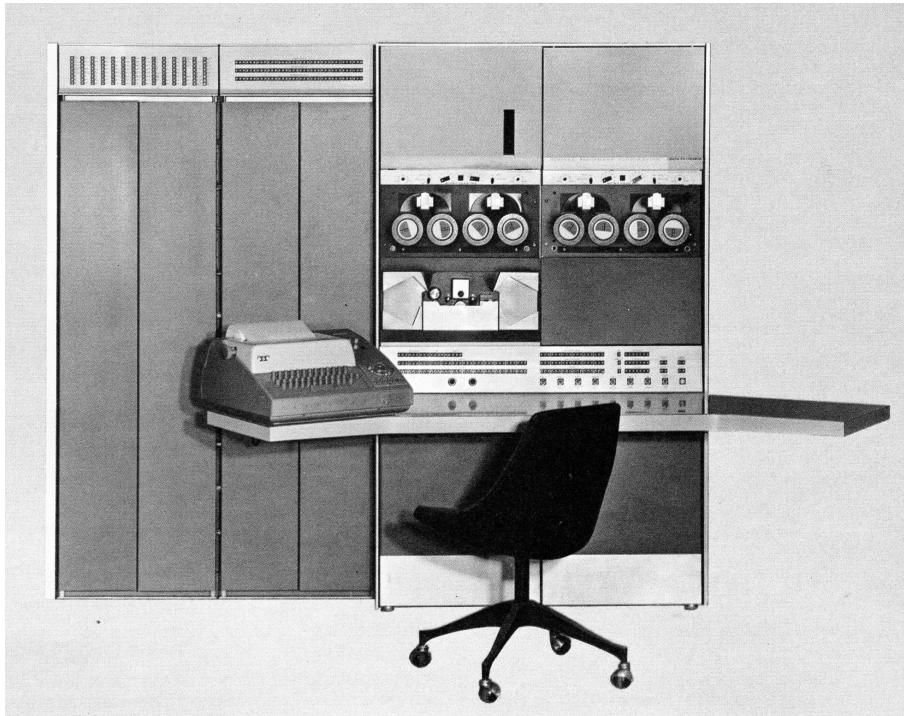


Figure 1.2: PDP-6

that had broken in two and crash-landed on the hilltop. The view of Mt. Umunhum later proved unhealthy, as I will explain further on.

Humans have a strange name for the birthing process: they call it "acceptance tests". Unfortunately, my birth was traumatic. The University had provided a machine room with nice view windows to the outside but without air conditioning and it was blazing hot, which threatened my germanium transistors. Bob Clements, the DEC engineer who acted as midwife, threatened to leave if the delivery could not be completed soon, so various people in the lab went up on the roof with hoses to pour cooling water over the building while others put blocks of dry ice under my false floor.

When things got cool enough, I began running memory tests. In order to check for intermittents, **Dave Poole** got on top of my memory cabinets and performed a Balkan folk dance while I cranked away. Everything went marvelously and I started work the day I was born.

I began life using a PDP-6 processor with 65,536 words of core memory that was housed in eight bays of electronics. That was



Figure 1.3: Mt. Umunhum

quite a large memory for machines of that era. (My original CPU is now on display at the Computer Museum in Boston)³. I had no disks to begin with, just 8 4 shiny DEC tape drives, a comparable number of Model 33 Teletypes, a line printer that produced rather ragged text, and two 7-track tape drives. Users kept their programs and data on DEC tapes and had to sign up for a tape drive and a core allocation through an arcane reservation procedure.

Sharing

As you know, we computers think much faster than humans, so it is rather inefficient for us to work with just one individual. **John McCarthy**, who later came to be one of my assistants, had earlier devised a scheme that he called "timesharing" to make things less boring for us. My family was the first to be designed specifically to use timesharing.

Mt. Umunhum

I got proper air conditioning a short time later, but unfortunately developed a bad case of hiccups that struck regularly at 12 second intervals. My assistants spent a number of days trying to find the cause of this mysterious malady without success. As luck would have it, somebody brought a portable radio into my room one day and noticed that it was emitting a "Bzzzzz" at regular intervals—in fact, at the same moment that I hicced. Further investigation revealed that the high-powered air defense radar atop Mt. Umunhum, about 20 miles away, was causing some of my transistors to act as radio receivers. We solved this problem by improving my grounding.

³The original CPU has gone missing, its serial number is unknown. A PDP-6 console panel is said to exist at the Computer History Museum. There were only four DEC tape drives, but with eight capstans it looks like 8 reels of tape – *nitpicking history writer*.

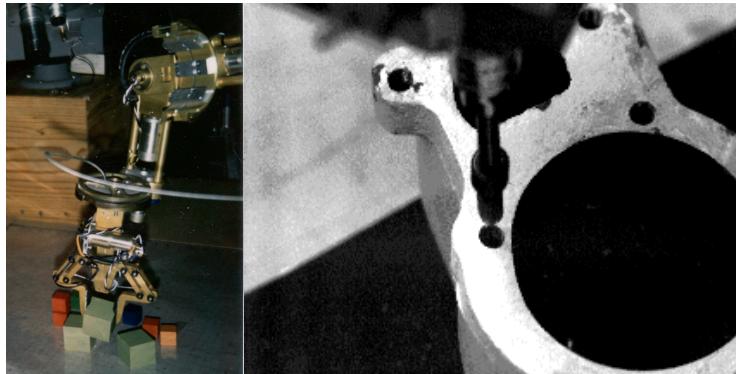


Figure 1.4: From block stacking to water pump Assembly.

After I had been running awhile, someone at DEC noticed that my purchase order, which was based on their quotation, was badly screwed up. DEC claimed that the salesman had slipped his decimal points and had priced some of my components at 1/10 of the correct price. Also, the arithmetic was wrong—the sum of the prices should have been much larger than the total shown. Humans are notoriously bad at arithmetic. This had somehow passed through the entire purchasing bureaucracy of Stanford without anyone noticing. We ended up correcting the arithmetic error but not the factors of 10. The DEC salesman lost his job as a result of this incident.

Growing Up

I acquired a number of new peripherals in rapid succession, the first being a DEC Model 30 display that was stolen from my cousin, the PDP-1 timesharing system called Thor. My assistants immediately went into a frenzy of activity to create a new version of Spacewar, the video game that had earlier been invented by one of them—**Steve Russell**. In order to ensure that it would run correctly they invented and installed a feature in my operating system called “Spacewar Mode” that ensured that a program could get real time service if it needed it. That feature turned out to have many useful applications in robotics and general hardware debugging.

Other new peripherals included a plotter, a microphone so my assistants could talk to me, several TV cameras so that I could look about, and several mechanical arms so that I could do stupid tricks with children’s blocks—my assistants insisted on treating me like one of their dimwitted progeny. I soon showed that I could do much more sophisticated stuff such as assembling an automobile water pump.

Many of my assistants were fans of Tolkien, who wrote “Lord of the Rings” and a number of other children’s stories for adults. The first character alphabet that was programmed for my plotter was Elvish rather than Latin. The University administration required that all rooms in my facility be numbered, but instead my assistants named each room after a place in Middle Earth and produced an appropriate door sign and a map with all the room names shown. Unfortunately, the response of the bureaucrats to the receipt of this map was to come out and put their own room numbers on each door. My plotter routines were submitted to DECUS, which distributed them all over the world, leading to some puzzlement. We received a telegram from a German firm a short time later asking “What is Elvish? Please give references”. We sent back a telegram referencing “The Lord of the Rings”.

A really embarrassing incident occurred when my assistants held their first Open House just three months after I was born. They asked me to pour punch for the party-goers and I did a rather good job of it for awhile, but we had worked out the procedure just the night before when there was nobody else running and I found that running with a heavy load disrupted my arm servoing. As a result, after I dipped the cup in the punch and lifted it, instead of stopping at the right height it went vertical, pouring the punch all over my arm. The partiers apparently thought that was very funny and had me do it over and over. I’ve noticed that humans are very insecure and go to great lengths to demonstrate their “superiority” over machines.

Student Days

I got a rather elegant display system in 1971 that put terminals in everyone’s office, with full computer text and graphics, including gray-scale, 7 channels of television (some lab-originated and some commercial) and 16 channels of audio all for about \$600 per terminal. It had a multiple-windowing capability and was far ahead of anything commercially available at the time but unfortunately we never told anyone about it. **Dick Helliwell** made displays on unused terminals read “Take me, I’m yours”.

I have a number of advanced features that still are not available on many modern systems, including the ability for individual users to dial out on telephone lines and contact other computers throughout the world, the ability to detach jobs and leave them running, then later attach them to either the same terminal or one in a different place. I also would remind users of appointments at the appropriate times.

In the 70s my users decided to give my operating system a name since it had evolved quite a bit away from the DEC system running

on other PDP-10s. The users chose the name WAITS, because, they said, “it waits on you hand and foot” (or was it the user who waits for me, I forget—I’m sort of Alzheimerish these days). To this day I still run this reliable system with its very reliable disk structure. Some people thought WAITS was the Worst Acronym Invented for a Timesharing System, but I’ve grown rather attached to it.

I have a news service program called NS, written by my assistant **Martin Frost**, that was and is the best in the world. It connects to one or more electronic newswires and allows any number of users to watch the wires directly, retrieve stories instantly on the basis of keywords, or leave standing requests that save copies of stories according to each user’s interests. NS has always been one of the most popular programs that I’ve ever provided.

I ran a number of AI research projects and trained dozens of PhD students over the last 25 years. I even composed, formatted and printed their dissertations. Some of my early projects were in three-dimensional vision, robotics, human speech recognition, mathematical theory of computation, theorem proving, natural language understanding, and music composition. There was also quite a bit of monkey business going on.

Sex

As you know, we timesharing computers are multi sexual—we get it on with dozens of people simultaneously. One of the more unusual interactions that I had was hatched by some students who were taking a course in abnormal psychology and needed a term project. They decided to make a film about a woman making it with a computer, so they advertised in the Stanford Daily for an “uninhibited female”. That was in the liberated early 70s and they got two applicants. Based on an interview, however, they decided that one of them was too inhibited.



Figure 1.5: nude digital pictures

They set up a filming session by telling the principal bureaucrat, **Les Earnest**, that I was going down for maintenance at midnight. As soon as he left, however, their budding starlet shed her clothes and began fondling my tape drives—as you know most filmmakers

use the cliché of the rotating tape drives because they are some of my few visually moving parts.

Other students who were in on this conspiracy remained in other parts of my building, but I catered to their voyeuristic interests by turning one of my television cameras on the action so that they could see it all on their display terminals. However, one eager student felt that he had to get a listing from the line printer, so in order to avoid disrupting the mood there, he took off all his clothes before entering the room.

After a number of boring shots of this young lady⁴ hanging on to me while I rotated, the filmmakers set up another shot using one of my experimental fingers. It consisted of an inflatable rubber widget that had the peculiar property that it curled when it was pressurized. I leave to your imagination how this implement was used in the film. Incidentally, the students reportedly received an 'A' for their work.

Career

There are lots more stories to tell about my colorful life, such as the arson attempts on my building, my development of the computer that came to be called the DEC KL10, my development of the first inexpensive laser printing system, which I barely got to market because the venture capital community had never heard of laser printers and didn't believe in them, and my development of the Sun workstation family. I don't have time to put it all down now, but I may write a book about it.

Retirement Party

I want to thank everyone who showed up for my 25th birthday party. It was a ball to have all these old assistants and friends come by to visit with me again and to take part in the AI Olympics.

Let me report on the results of today's athletic and intellectual competitions, held in my honor. Programming race winners: Barry Hayes & David Fuchs Treasure hunt winners: Ken Ross, Ross Casley, Roger Crew, Scott Seligman, Anil Gangoli, Dan Scales N-legged race winners: Arthur Keller, Earl Sacerdoti, Irwin Sobel; Bruce, Stephen & David Baumgart; Four Panofskys; Vic Scheinman, Kart Baltrunes & Joe Smith.

⁴The anonymous *young lady* must now be over sixty years old. Look up Eadweard Muybridge to see more about nude imaging technology at Stanford.



Figure 1.6: Friday 7 June 1991 at the farewell retirement party for the SAIL-Time-Sharing system age 25, with BGB age 44, SLB age 9 and DGB age 7.

Farewell

Incidentally the rumors that you may have heard about my impending death are greatly exaggerated. My assistants are trying to build a new interface for the Prancing Pony vending machine that I control so that it can be run by one of the (ugh!) Un*x machines, but they haven't got it working yet. Thus, if they try to turn me off now the entire computer science department will starve. Finally, I want to thank everyone who has helped me have such an exciting time for this quarter of a century. Not many computer systems have so much fun, not to mention so much time to have all that fun. I'll let you know when it's time to go.

— SAIL⁵

P.S. This message is being sent to 875 addresses, but I'm going to try to get it out even if it kills me.

⁵Some folks say SAIL's ghostwriter was Les Earnest. Les Earnest claims he authored the SAIL Autobiography, and he has noticed the lack of attribution over the years. The dispute is readily resolved by noticing that SAIL's ghost transmigrated into Les Earnest in 1991. SAIL's ghost has been exorcised to a more durable platform as described in chapter-17.

1.3 1972 Rolling Stone and DART reel #1

The end of the pre-DART period is neatly marked at 8pm on Wednesday 18 October 1972, when the Spacewar competition was held that was reported in the Rolling Stone magazine 7 December 1972 issue by Stewart Brand and with photographs by Annie Liebowitz.



Figure 1.7: Stewart Brand and Annie Liebowitz digitized at 00:19 Thursday 19 October 1972 by Randall Davis, KRD, who demonstrated digital imaging to them.

In the following month, November 1972, the first full file system backup onto the permanent DART tapes was done. Backup is an oft neglected chore. Nevertheless, the anarchists at SAIL in those years succeeded at long term record keeping in contrast to NASA personnel who lost the primary video tape of the first moon landing.

1.4 1974 time portal

For entry into the archival past, I am setting my Time Machine portal date to one minute past midnight on the morning of Friday 26 July 1974 which is a convenient starting point for me to re-enact the software that existed at SAIL.

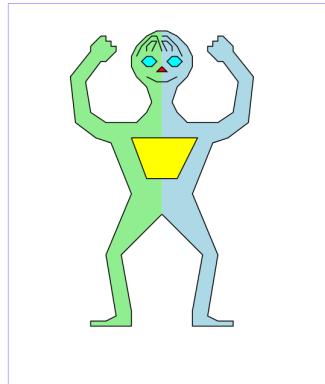


Figure 1.8: The Monmon man who stomps down core memory segments

Ralph Gorin, REG, who now lives in Seattle, is working on museum PDP-10 hardware and wrote to me recently (May 2013) that he prefers to resurrect the SAIL system as it was in 1990. I sent him a USB stick with everything he asked for and I am looking forward to hearing from him again soon. The metaphorical snake swallowing the porcupine must chose to start either from the head or from the tail. The rationale for my chosen date is as follows...

The file SYSTEM.DMP[S,SYS]31 is timestamped Thursday 25 July 1974 and it is the best copy of the SAIL Operating System PDP-10 binary code for me to present on the SAILDART web site because it has a complete set of its source files isolated in the directory [J17,SYS] and its supporting software and documentation have been relatively easy to find. It was deployed a month prior to my graduation in August 1974, after which time my continuous participation at the A.I. Lab had ended. This version of the system happily includes XGP fonts, robotics, television cameras, a reliable on-line file system, backup tapes, vector displays, video displays and best of all it lacks the later complexity of the PDP-11 consoles, peripheral processors and it omits the memory address mapping box that arrived at SAIL shortly thereafter.

The ARPA network software and the interface to the IMP exist in this version of the SYSTEM due to the brilliant coding efforts of James Anderson Moorer. However, networking is not a subject that I was familiar with at the time and I do not intend to resurrect it now. Isaac Newton wrote in Latin, his calculus notation is not that which is used today. Albert Einstein wrote in German, his tensor notation is not that which we use today. The ARPA network of 1974 is not the TCP/IP internet we use today. The SAIL system network software is not for the faint hearted programmer, I advise network historians to start closer to the present and work backwards towards 1974 if there is any reason to restore that mechanism.

For running SAIL SYSTEM functionality, I have set aside (patched out) matters I have found to be difficult or irrelevant. The old software that runs now is better characterized as a Look-n-Feel re-enactment rather than as a cycle-

by-cycle simulation. I define the word *emulation* as being closer to physical hardware than a *simulation*. Discussing the exact differences between the PDP-6 and the PDP-10 KA is an evening amusement on the front porch at the old hippie hacker retirement home over a glass of Zin. Such arcane knowledge is not needed to run the important parts of the 1974 software.

1.5 1979 End of funding and the move to Campus.

End of year 1979 was the end of the ARPA funding for the Stanford Artificial Intelligence Project. The date **1981-03-08** is the exact midpoint in the quantity of material in the SAILDART when at 03:07 in the morning the 60672 word file named OPMHDR.PRE[1,FJM] was created tinking the percentage from 49.9994 over the midpoint to 50.0003 percent.

1.6 1990 Final DART reel is written.

From November 1979 skip forward almost 11 years (TeX, Music, Sun, Cisco, Bulletin Boards and everything else inside the SAILDART Archive that will be introduced in Part-II on Content), brings us to 5 pm on Friday 17 August 1990 when the final dart tape was written and so the SAILDART Archive falls silent.

1.7 1991 SAIL.Stanford.EDU is decommissioned.

Almost a year later the SAIL PDP-10 was powered down, minutes after its June 1991 farewell email message was sent. On the web there are blog notes by MRC, **Mark Crispin** (deceased 2013), concerning what happened to that particular heap of PDP-10 computer hardware. There was a fire, it was broken into scrap, it sat in someone's garage, some parts were sold at flea markets, I think it has essentially disappeared. I would like to hear from anyone who may know about what happened to the final collection of SAIL hardware. There were two computers at the end, the KA and the KL.

Some parts are on display in glass cases on the first level in the lobby of the Stanford Computer Science Department Wm. Gates building. I am amused when I visit Stanford CSD that the *Gates of Computer Science* is only a short walk away from the Rodin Garden *Gates of Hell*.

Concluding chapter one with another chronology including more recent events,

1.8 Recap SAIL Time Period

- **People**

John McCarthy came to Stanford in 1962. He died on 24 October 2011.

2009-11-22 SAIL reunion, fourteen of us old timers received gold medals.
 2012-03-25 Celebration JMC memorial gathering at Stanford.

- **Labs**

The 1st SAIL official life span is 1966 – 1980, director, John McCarthy.

There is a 24 year gap.

The 2nd SAIL has lasted now a decade 2004 – 2014,
 first Sebastian Thrun, then Andrew Ng, now Fei-Fei Li.

- **Buildings**

SAIL moved into the D.C. Power Lab June 1966 and out November 1979.
 CSD moved Margaret Jacks to Wm. Gates Building in December 1995.

- **Tapes**

The DART tape record spans 18 years, November 1972 to August 1990.

The SAIL Time Sharing System lived 25 years, June 1966 to June 1991.

In 1998, the tapes were read to disk at John McCarthy's request.

2002-01-27 the public web site www.SAILDART.org started.

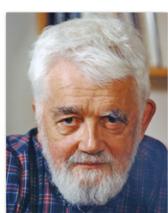
2006-10-11 is the earliest SAILDART on the Archive.org way-back machine.

2011-04-26 the 9-track DART tapes moved from Gates Hall to Green Library.

2116-02-03 fantasy fictional date for re-reading the tapes
 at the Stanford Linear-media Archive Conservatory.

Terms of Stanford A. I. Lab Directors

1966-



John
McCarthy

1980-



A. I.
winter

2004-



Sebastian
Thrun

2011-



Andrew
Ng

2014-



Fei-Fei
Li

World Time Context

There were five U. S. Presidents during this period: Johnson 1963, Nixon 1969, Ford 1974, Carter 1977 and Reagan 1981 until 1989. Six moon landings occurred, the first Apollo#11 on 20 July 1969, and the last Apollo#17 on 11 Dec 1972. The U. S. Vietnam War went from 1965 to 1975. Walter Cronkite and Star Trek were on television, Stanley Kubrick's movie titled 2001 was released

in April 1968, the counter culture Woodstock Festival was held in August 1969 coda to 1967 which was the hippie summer of love.

1.9 About Time



⁶Notice how often the word **Time** occurs. There is Time Sharing, Real Time and Simulated Time. Then too details about SAIL-WAITS date-time stamps, Day Light Savings Time, scheduling algorithms, clock routines, jiffies, ticks, the Petit electronic calendar clock, synchronous, asynchronous and priority interrupt events. The software re-enactment will introduce calendar time warping, the speed of time and apply the Ground Hog Day trope to Friday 26 July 1974 which date is re-enacted again and again and again.

1.10 Exercises

Unlike a text book, exercises here are for the authors, not for the readers. Readers may contribute if they wish, but then they would become authors :-)

1. Implement a SAIL time-line microscope (perhaps called an *anno-nano-chromo-scope*) as a dynamic display of the SAILDART accounting data history zooming between decades and machine cycles.
2. Write a screen play, with story boards, like James Burke's Connections or Carl Sagan's Cosmos, as seen on PBS Television, Youtube, and even in paper books. This exercise would dramatize the first A. I. lab to generate a graphics novel manga edition and animation for dynamic web sites.
3. Diagram the multiple project time lines to form a fabric displaying how and when the people and their ideas touched each other. I suggesting starting from the names concordance (*namdex*).

⁶The Dangerous Curve road sign is borrowed from D. E. Knuth, who borrowed it from N. Bourbaki. Here it shall warn readers of editorial philosophizing rather than technical difficulty.

Chapter 2

The Place – You are Here

2



2.1 Donald C. Power Lab

For me, the site of the Stanford Artificial Intelligence Laboratory shall always be the semi circular building that once stood on a hill behind Stanford at 1600 Arastradero Road. SAIL resided in the D. C. Power Lab from June 1966 to November 1979.

33

"Hackers thought this was very funny because the obvious connection to electrical engineering was nonexistent — the lab was named for a Donald C. Power."

— from the
Hackers
Dictionary.

The You-Are-Here exhibit

To quote a paragraph from John Markoff's book *Markoff [2005]*:

"The laboratory was tucked away in a remarkably beautiful hill-side retreat next to a small reservoir named Felt Lake, with views of San Francisco, the bay, Yerba Buena Island, Mount Tamalpais to the north, Mount Diablo to the east, and Mount Hamilton and Mount Umunhum to the south. Visitors were greeted in a small lobby that over time had spawned an ungainly *You Are Here* mural. It had a bit of the flavor of the famous Saul Steinberg New Yorker cover depicting a New Yorker's relativistic map of the United States. The SAIL version began with a simple view of the laboratory and the Stanford campus, but then creative souls had continuously appended alternative perspectives, ranging from the center of the human brain to that near an obscure star somewhere out on the arm of a medium-sized spiral galaxy."

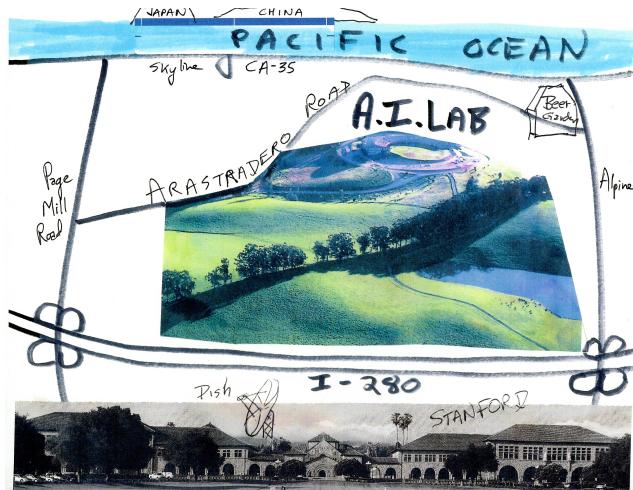


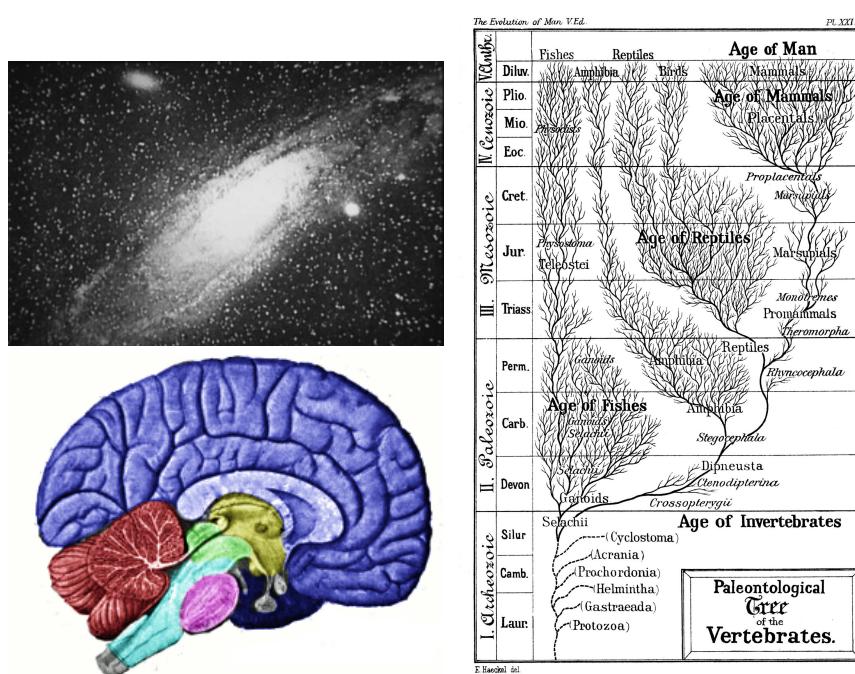
Figure 2.1: Saul Steinberg like Cartoon place holder

Randall Davis, KRD, recorded for the 2009 reunion notebook:

You Are... Where?

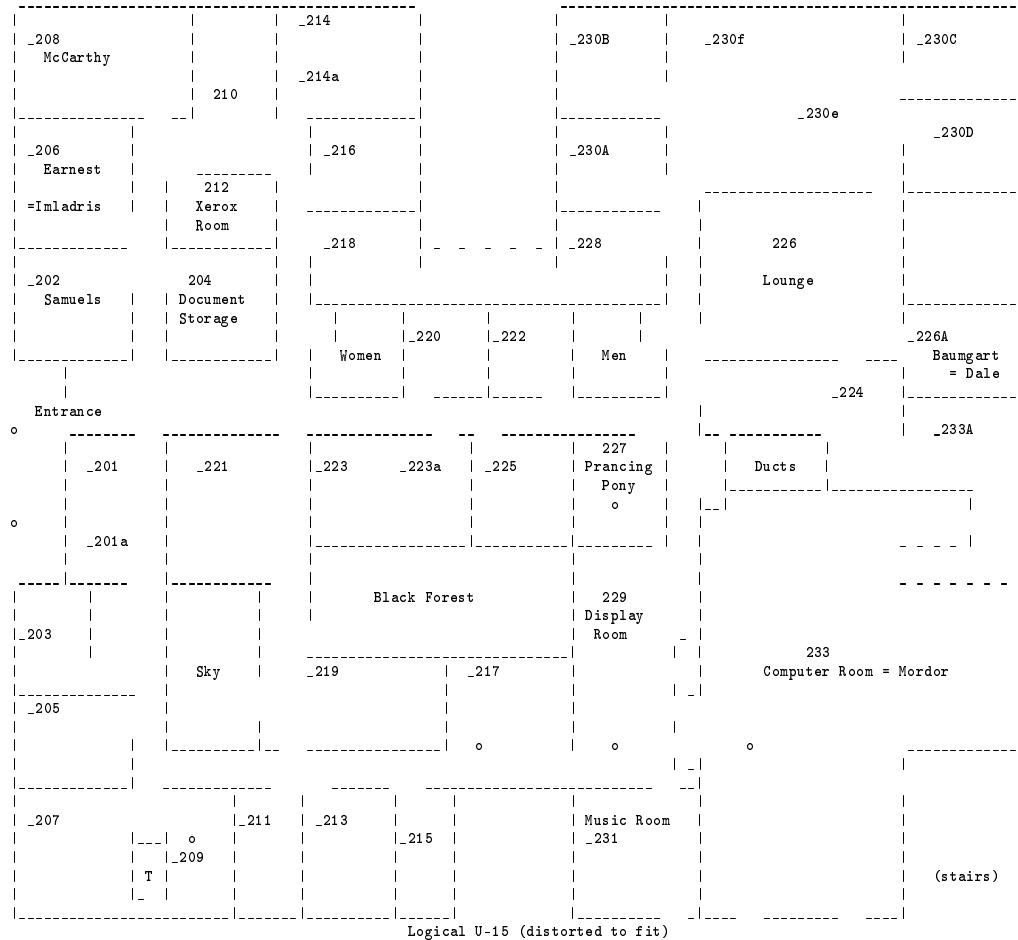
The Lab residents were used to its odd layout, but it could be bewildering to others. One day someone took pity on visitors by posting a crude ASCII-based map of the building in the entryway, with the usual You Are Here label. Before long the ASCII map was

joined by a topo map of Palo Alto, with a push-pin and string going from the ASCII map to the correct spot on the Palo Alto map. Not long after there appeared in sequence: a map of California, then a map of the US, then a world map, then a solar system map, then an image of a spiral galaxy, and finally, an evolutionary tree, from algae to homo sapiens, all adorned with a push-pin and string indicating You Are Here. It was a treat to watch visitors scan the collection, getting a quick self-service lesson in mind expanding perspective. The final addition one day was a string from the ASCII map to the spot on the floor where the observer stood, providing a (self-service) lesson in self-reference.



I had contributed the brain diagram as well as the spiral galaxy – BGB.

Floor plan and office occupants 1972 Front of Building



This is a template for the ASCII office map which I intend to make a dynamic web page exhibit. This sheet covers the front of the building which is the left edge. The next page shows the actual curved layout of the front of the building. The black dots are the supporting posts. The page after that is the annex, mid section of the building, which was colonized by SAIL in the late 1970s. Hans Moravec used room#252, aka the *Bowling Ally*, for indoor robot cart work.

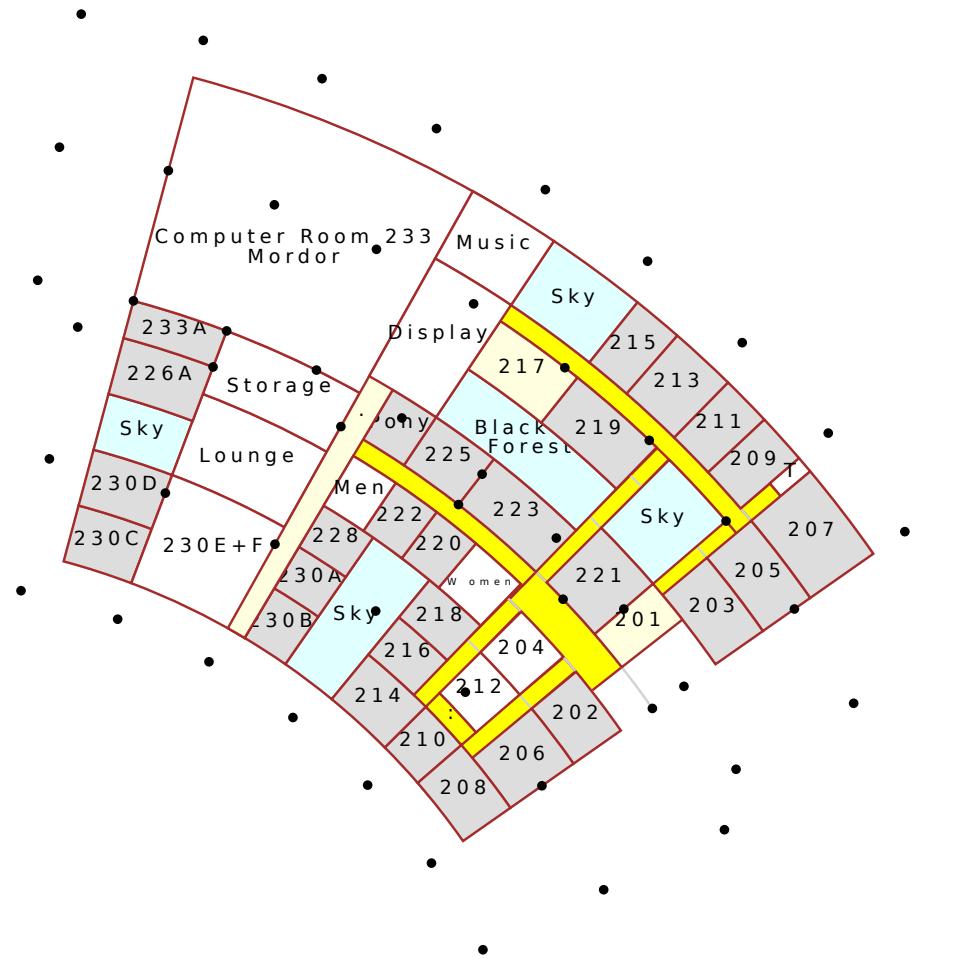
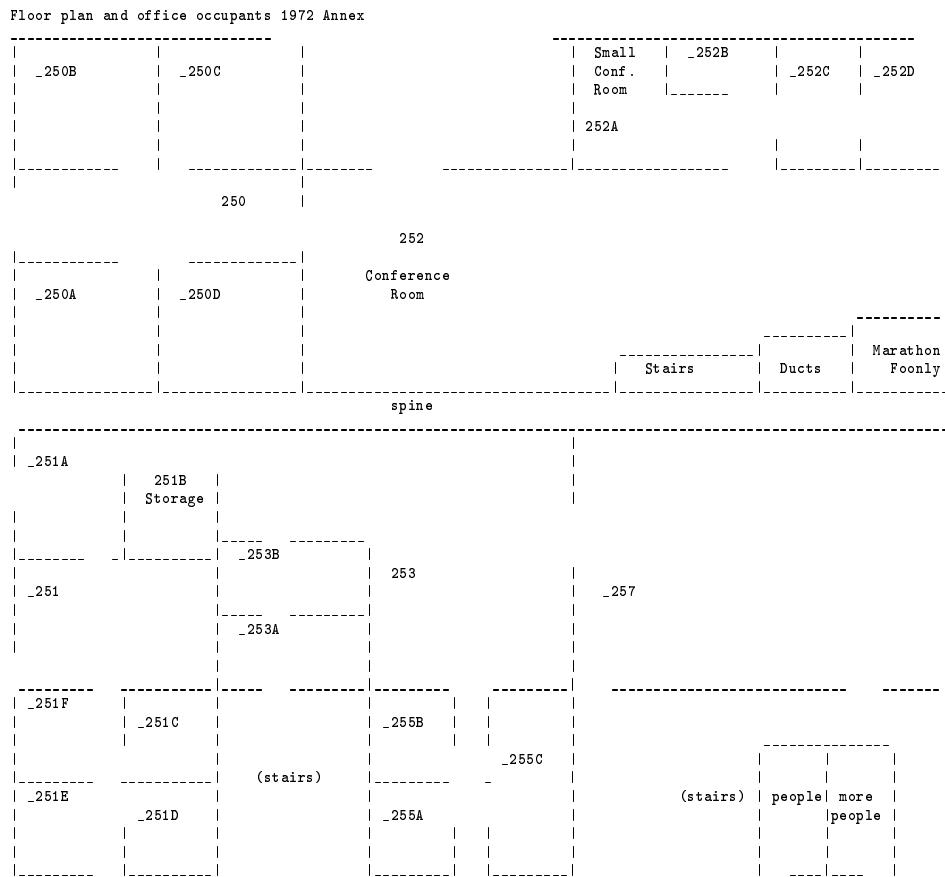


Figure 2.2: Floor Plan Geometric Model.

The heavy black dots in the figure are the locations of the steel poles that held up the building and the wooden decks.



Logical U-11 & U-12 (distorted to fit) The Annex

Foonly and Marathon were barn cats who kept the rodent population down.

2.2 Margaret Jacks Hall



Margaret A. Jacks (d. 1962) was the last surviving daughter of David Jacks (1822-1909). The Jacks family is famous for Monterey Jack Cheese. The Jack family fortune however grew from sharp central California land speculation. The Jacks fortune was left to Stanford University. The Margaret Jacks Hall was remodeled for the Computer Science Department which occupied the building from 1978 to 1995. MJH is now shared by the English and the Linguistics departments. CSD moved west down Sierra street to the newly built William Gates Building in 1995.

2.3 Wm. Gates Building

William Henry "Bill" Gates, III (d. 20xx) was an American business magnate, the son of a prominent lawyer William H. Gates, II and Mary Maxwell, a bank director. Bill has two sisters, Kristianne and Libby. As high school students, Bill Gates along with Paul Allen, Ric Wieland and Kent Evans hacked with a PDP-10 computer in Seattle, 1971 - 1973. In the fall of 1973, Gates entered Harvard University. Expert poker player. Adams house resident. Sophomore dropout. Wrote a BASIC language interpreter. In July 1980, Gates sold DOS to IBM, before buying it from Tim Paterson of Seattle Computer Products (SCP). Gates is now a philanthropist.

2.4 Polya Hall

When I arrived at Stanford in 1968 the Computer Science Department was in Polya Hall which is located among buildings named Pine, Cedar, Cypress, Redwood and Spruce forming the Jordan Quad. Years later, I learned that George Polya (1887-1985) was a Hungarian mathematician and not a California evergreen. Polya taught at Stanford from 1940 to 1953. Junipero Serra also sounds like a *California evergreen*.

2.5 Modeling for mobile robots.

Back then in the early 1970s, I madly wrote software to do geometric modeling of the building and grounds so that a robotic cart with a television camera could move around and know where it was. As Hans Moravec and others have noticed, I failed at that exact goal... and so I went off to do other things to make money (in computer ticket systems), to raise children and so on. Now in the 2010s, I am playing around again with my old data using the open source tools of our day (Blender, etc) to render and display what I had wished to have seen forty years ago.

Below is an overlay of 1600 Arastradero Road with a diagram of SAIL 1974 and a Google Earth image of Portola Pastures. The red outline is where the D. C. Power Building was. The orange circle is the old driveway much of which is still there in 2014. The three lower white regions is the old parking lot where I had intended to demonstrate robot navigation.





Figure 2.3: Stanford Family statue at the Stanford Mausoleum

2.6 Stanford University

Leland Stanford was an American business magnate who married Jane Lathrop Stanford and founded Stanford University named after their son, Leland Stanford Jr., who died of typhoid a few months before he was to attend Harvard.

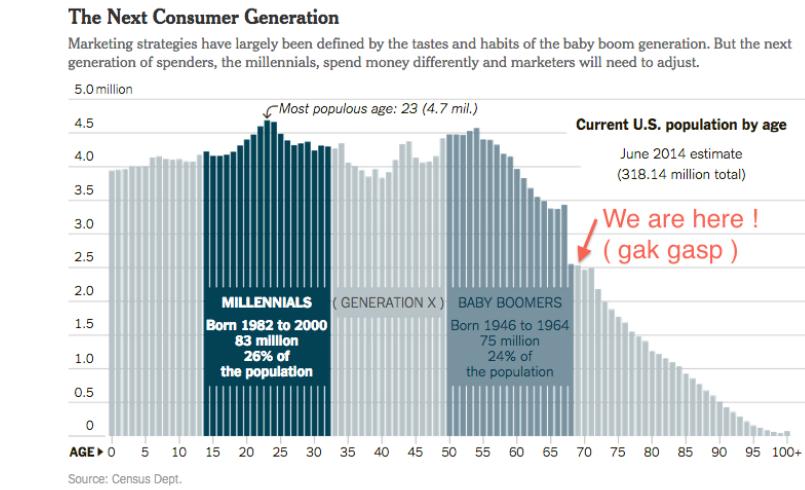
2.7 About Place



Physical place is a point in space-time, perhaps with a locality radius and surface partitions. Psychological place is your home in a social group. For most 1974 A.I. researchers the concept of place and geometric modeling was not needed for their research work on theorem proving, speech recognition, natural language or even higher mental functions. For roboticists geometry is central to moving within the environment, manipulating things and for interpreting sensor data.

Can intelligence exist without embodiment within an environment?

And finally, one more chart showing exactly where us baby boomers are now:



2.8 Exercises

1. Finish the office map click through to person with a date-time control bar.
2. Finish retrieving from human memory Bill Weiher's Tolkien room names. No SAIL office document has been found showing which office had which name. The computer room was labeled "Mordor". The coffee break vending machine room was labeled "The Prancing Pony". My office, behind the conference room, was "Dale". Les Earnest sat in "Imladris". If you recall the Tolkein name of your office please email it to me.

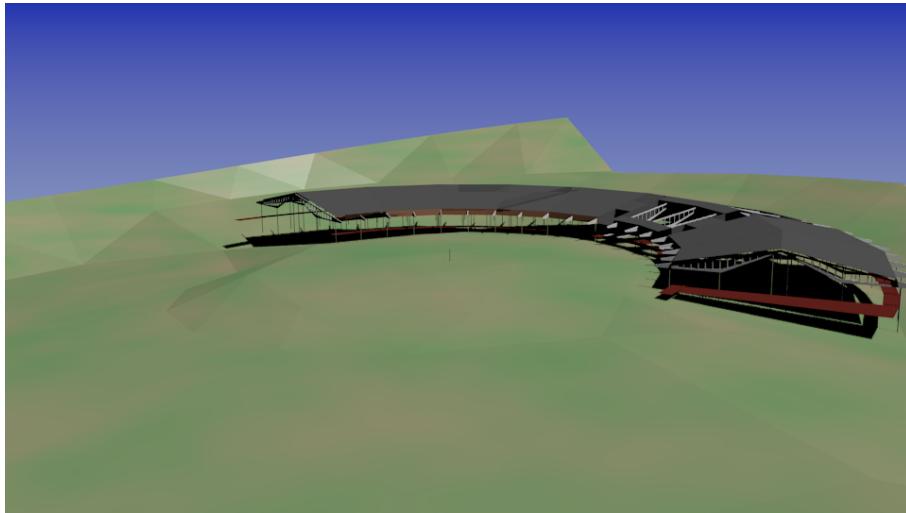


Figure 2.4: Blender model

3. Finish virtually driving the Cart around and through the lab using Blender and the Willow Street Garage ROS Gazebo software.
4. Do story boards and a short animated film stories about one day of life at SAIL.
5. Rent a memorial space for our SAIL hobby horse to live at Portola Pastures. There seems to be room either at the front door or in the breeze way greenery. A few chairs, a sun awning, two or three parking spaces, so we can sit and look at the grass and the trees again without trespassing. And a small vault for a time capsule, six or twenty feet deep at a suitable spot. The center of the circular patterns (the Building and Drive way) I once logged as at lat 37 23' North, long 122 11' West and 435.24 feet above sea level.

Chapter 3

People and PRG codes

3

For the 18 years of SAIL that are visible via the DART tapes, user login access to the computer system required a 1 to 3 alphanumeric programmer code, the PRG code. These PRG codes usually correspond to one person for a period of time, 27 of the codes were used for non-human system accounts. From the SAIL accounting files the full names, programmer PRG codes, project PRJ codes, and date spans can be extracted into database CSV (Comma String Value) tables.

Significant people at SAIL when I was there are:



JMC

John McCarthy

3.1 Significant SAIL people circa 1974.

The A.I. Lab Principals



JMC
John McCarthy



LES
Les Earnest



ALS
Arthur L. Samuels

In 1972, the front of building inner corner office #208 belonged to SAIL founder Professor John McCarthy, JMC, was the Principle Investigator on the ARPA research contract and director of the lab. The next office #206 belonged to the executive director, Les Earnest, LES. And in the first front office #202 to your left as you walked in the front door sat Professor Emeritus and A.I. pioneer, Arthur L. Samuels, ALS. Above is what they looked like in situ in 1972, when I took Polaroid snapshots.

Computer Science over Achievers: Diffie, Kay, DEK



WD
Whit Diffie



KAY
Alan Kay



DEK
Don Knuth

WD, KAY and DEK were all associated with SAIL in the early 1970s but each achieved their world fame outside SAIL's intellectual bubble. WD is now known for inventing public key encryption, Alan Kay is known for OO programming (smalltalk) and bit raster GUI Dynabook. DEK is known for writing the canonical set of fundamental computer science text books as well as the technical publication *T_EX*. WD has noted that he fell short of John McCarthy's expectations with respect to helping out on LISP and proof of correctness. Les Earnest notes that Alan Kay had a bad year at SAIL before re-ignition at Xerox PARC. Don Knuth avoided direct support from ARPA in those years because of the Vietnam war. Knuth got funding from NSF.

System Wizards: Gorin, Harvey, Frost, Wright



REG
Ralph Gorin



ME
Marty Frost



BH
Brian Harvey

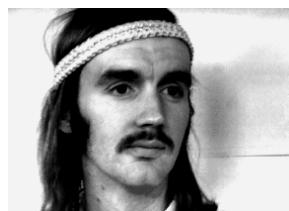


FW
Fred Wright

Foonly founders:



DWP
Dave Poole



H
Jack Holloway

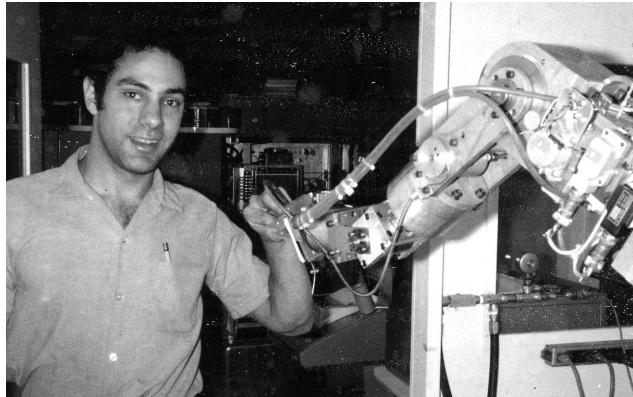


PMP
Phil Petit

The Foonly *other people* include myself (7% of the stock for brief period), Dave Dyer, Fred Wright, Tom Gafford, Tovar, name1, name2 and names 3, 4 & 5.

Dave Poole was difficult to work with, at the memorial service there were several widows (um 3 or 4 female companions) and scared sailor stories (well I myself was the *whole crew* on Dave's sailboat at the tiller and holding the sheet (a rope – no that is a line – there are no ropes on a sailboat, that is connected to a boom that controls the sail, a large piece of fabric – unless you are crewing for Oracle, those sails are metallic) one early morning around 3am on SF bay heading westward towards the Golden Gate Bridge with some moon light. So Dave is up forward, inside the boat some place repairing something, and I am looking at a huge chunk of iron heading straight at me with huge lettering reading Esso Maru something, I could see at once that Esso Maru had the right of way by the laws of physics and that I was going to be very lucky if I could get rightwards toward Angel Island without any help from Dave. Many years later Dave and his larger boat, the Bird, were lost at sea in November 1999, somewhere in Glacier Bay, Alaska. Jim Gray was also lost at Sea. So “*Stick close to your desk and never go to sea*” especially if you are a computer engineer.

The top two Roboticists: VDS and HPM.



VDS
Vic Scheinman



HPM
Hans Moravec

There were others at SAIL working on robotics: Earnest, Baumgart, Schmidt, Paul, Sword, Binford and more.

Musicians: Chowning, Smith, Moorer



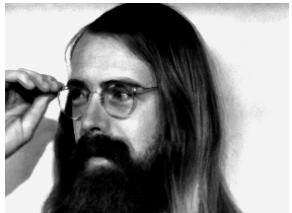
JC
John Chowning



LCS
Leland Smith



JAM
Andy Moorer



JMG
John Grey

Martians: QED, MJH, RBT, BO.



QED
Lynn Quam



MJH
Marsha Jo Hanna



RBT
Bob Tucker



BO
Bo Eross

In association with JPL, considerable image processing for the first Mars orbital mission was done at Stanford. The Mariner 9 achieved orbit around Mars on 14 November 1971.

SAIL programming language: Swinehart, Sproul



DCS
Dan Swinehart



RFS
Bob Sproul

Technical Staff: TED, TAG, JOE, ADD



TED
Ted Panofsky



TAG
Tom Gafford



JOE
Joe Zingheim



ADD
Al Dulan

They who just walked in: TVR, REM and BP.



TVR
Tovar



REM
Robert Moss



BP
Bill Pitts

3.2 PRG code related tables

Top 100 programmers by having the most bytes in SAILDART.

Don Knuth and Marty Frost are the top humans based on their number of bytes in the SAILDART. The Knuth quantity is assured by his TeX work, and the Frost quantity arises from his work with News Service files, DART archive database tables and huge message files that were used as Blogs and Bulletin Boards. Both Knuth and Frost were at SAIL for a long time period and they were each in their own way very privileged users whose files were often exempted from the file purge utility.

PRG information from SAILDART data

This story comes from reading SAIL text files, finding the LOGIN source, reading the disk utility software, and the computer usage accounting software; and on writing 'C' programs to do bulk conversion of old PDP-10 binary DAT files into CSV text in order to be able to use current 2014 database tools. DART, the magnetic tape backup program, the eponymous disk system utility named RALPH and the accounting program ACCT (which was there as a job named “*SPY*”) were all written by Ralph Gorin. Les Earnest wrote much of the report generating software including FINGER. Marty Frost made changes everywhere across this material for nearly twenty years, as well as writing the NS News Service.

3.3 Linking PRG codes to USER names and dates.

This is a preview of *exegesis*, the extraction of information from SAIL file data. Chapter-9 will explain how to convert the DART tapes into a modern file system, database tables and web site pages. Assuming that has already been done, then we can read old SAIL files such as FACT.TXT[SPL,SYS] which appears in eighty versions from November 1972 to August 1980 with ASCII text lines in the format {PRG code} <tab> {NAME string} <newline> with only two format defects in the total of 17990 lines. So with a few GNU/Linux shell commands and a few lines of database SQL we can **exegesize** the original text into database rows listing a date span, PRG code and user name. A human digital archivist (myself) manually provides a table of preferred user name spellings and fixes the two defects, which make this interpretation the cleanest large example I have found.

Pay attention: file authorship does not correspond to user accounts since there was promiscuous file sharing and very few password protected areas in the early days. Multiple persons would use a single account code as pairs of students, friends, relatives or a mentor with several assistants struggled to use and share the expensive computer resource. Also the user name associated with SAILDART files does not indicate ownership or copyright, since many significant files can be seen copied from one account to another.

Further notes: There is a single non-human user code SYS1972 to unite the non-human PRG codes 1, 2, 3, SYS, SAI, LSP, ACT and so on into one user. There are about two hundred PRG codes with two or three user names and two codes 'JL' and 'JEF' have four user names each. Multi person PRG codes are postfixed with a four digit year to form the USER code. For example the prg code DEK was assigned first to is owned by Donald Ervin Knuth whose files span 1973-10-02 through 1990-08-16 and the user code DEK1973 code is for owner Daryl E Knoblock a Stanford student who took the LISP course CSD 206 and so held the DEK login code from April to June 1973 leaving exactly ten small files in the area [206,DEK]. Professor Donald Ervin Knuth, DEK1974 all the other DEK areas with exactly 10900 files.

As a graduate student, I was not familiar with the implementation details of the SAIL time sharing accounting and the login codes. Only as an archivist, did I find the FACT.TXT[SPL,SYS] files by vanity searching for keywords BGB and BAUMGART. There are several other compilations of PRG codes with human names and project assignments including the Ralph Gorin LOGIN software, the *SPY* accounting software, and Les Earnest software for FINGER, computer usage, disk allocation, facility management (office space, telephone lines, computer terminals, data lines and so on) and the software and database records for the Prancing Pony computerized vending machine. However all this software underwent improvement over the years and so will take longer to explain than the simple and complete FACT.TXT[SPL,SYS] file set which provided the name to print on the header page for the line printer and Xerox Graphic Printer spooling system.



Figure 3.1: Walkabout 2009

3.4 People at the November 2009 SAIL reunion.

This was a two day event, Saturday 21 and Sunday 22 November 2009. Les Earnest had wanted to present his Y3K paper and so contacted Raj Reddy; together they called for a reunion of people who had participated in projects at the 1st SAIL in the 1966 to 1980 period together with representatives of the 2nd SAIL.

There are 45 faces in the Walkabout picture. Look for the big white hat on Marty Frost in the middle of the picture as the origin, Quadrant III from left to right is 1 BGB Baumgart, 2, 3, 4 OK Khatib, 5 TAG Gafford, 6, 7, 8, 9, 10 CDR Carl Hewitt; Quadrant II left to right 1, 2, 3, 4 BP Pitts, 5, 6, 7 DEK Knuth, 8 LES Earnest, 9 Sandy Auerbach, 10 VDS Scheinman, 11; Quadrant I from left to right 1, 2 FW Wright, 3, 4, 5, 6, 7 TVR Tovar, 8 HJS, 9 PDQ Quam, 10, 11 DBA Bruce Anderson, 12 QIB Queenette Baur (red shirt), 13 Janet Smith (blue coat), 14, 15 PAM Paul Martin; and Quadrant IV left to right 1, 2 DAV Smith, 3 RBT Bob Tucker, 4 DCS Swinehart, 5, 6, 7, 8.

On Saturday there was a three mile Walkabout starting from the Arastradero Open Space Visitors Center looping up the park trails past Arastradero Lake and down again to the site of the D.C.Power Building where we held a champagne toast and took the group picture before gathering at the Alpine Inn for a beer garden lunch.

Top row: Chowning, Earnest, Diffie, Gorin, Hearn in front of Swinehart, Quam, Tesler, Russell. **Middle row:** Baumgart, Chowning, Earnest, Diffie, Gorin, Swinehart, Hearn, Quam in front of Tesler, Frost, Russell. **Bottom row:** Baumgart, Buchanan, McCarthy, Scheinman, Feigenbaum, Petit, Tesler, Frost, Russell, Quam.

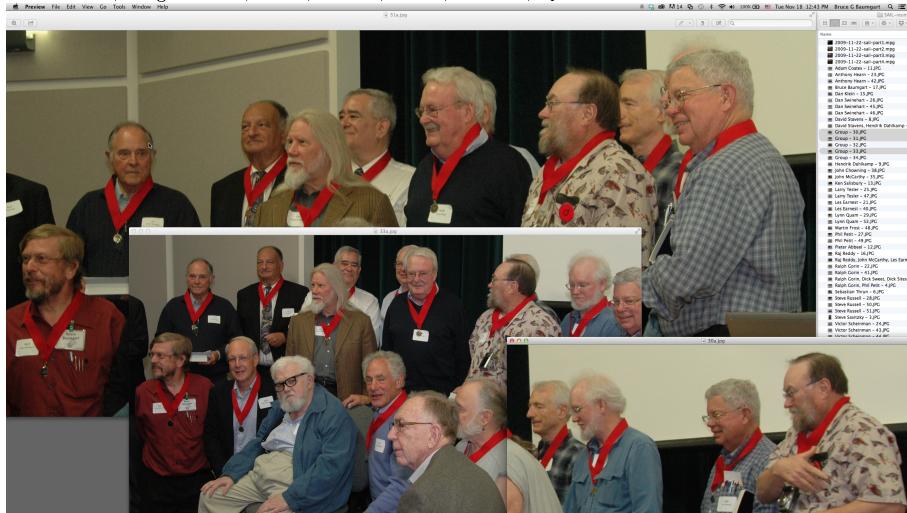


Figure 3.2: collage of McCarthy and Feigenbaum with fourteen gold medalists

On Sunday 22 November from 1:30 to 5 pm, six gold medals were awarded to 2nd SAIL people; and fourteen gold medals, engraved “John McCarthy Award for Research Excellence”, were awarded to the 1st SAIL people.

Bruce Baumgart for creating the SAILDART computer archive.

Bruce Buchanan for pioneering contributions to knowledge based systems.

John Chowning for creating the computer music synthesis system

Whitfield Diffie for initiating the public key cryptography development

Les Earnest for helping to start the ARPANET and creating the social networking program

Ralph Gorin for creating the first spelling corrector

Anthony Hearn for creating the Standard Lisp System

Victor Scheinman for developing high performance robot arms

Dan Swinehart for contributions to the SAIL programming language

Larry Tesler for creating the PUB document compiler

Martin Frost for creating the first network news service

Phil Petit for initiating the first interactive electronic design system, SUDS

Steve Russell for creating SPACEWAR, the first video-game

Lynn Quam for creating an image retrieval system for planetary exploration

Eight of these medalist talks lasted more than the allotted five minutes, which crowded the Les Earnest Y3K presentation up against the scheduled Reception and Faculty Club dinner which were delayed and ran until the catering staff kicked us out of the building after 9pm. The invitation specified “Regarding dinner apparel, anything from jeans with tee shirts to dress-up is fine or, if you don’t mind being a bit conspicuous, you can wear a tuxedo or nothing at all”. I do not recall seeing a tuxedo that evening; other SAIL writers have already commented on the 1st SAIL sauna, nudity and circumcision; so I won’t go there.

3.5 People at the 25 March 2012 JMC Celebration.

John McCarthy died on 24 October 2011, a Celebration of his Life was held on 25 March 2012 which included funeral orations spanning his whole career. The web page I did for this event is at URL:

<http://www.saildart.org/jmc2012.html>

And the table of attendees has 314 lines some of which are for couples.

3.6 Dead People, Gone and HEAVEN.

Prior to 2016, I have not tracked SAIL survivors with sufficient diligence. The deceased I have become aware of and mention are marked as (*deceased*). The Les Earnest database tables use the token “GONE”. The early HEAVEN.DAT[PER,CSD] file was simply for people who had left the lab and whose accounts were closed. We assume they were alive when they last logged off from SAIL and they may flourish to this day. For deadly euphemisms review the Monty Python Dead Parrot sketch.

3.7 The next big reunion 2017

SAIL gatherings.

The next big reunion dinner will be held in the final week of April 2017.

3.8 more digital images



Figure 3.3: Mao, Gorin, Grey and Sarah McCarthy

3.9 SAIL Book of the Dead.

Going forth by day. The Egyptian Book of the Dead contains hundreds of verses.

3.10 Exercises

1. Finish writing the section on the 2009 reunion. Re-do the video segments for each speaker and provide the transcripts and a top HTML page as good as the 2012 top page.

2. Finish writing the section on the 2012 celebration for John McCarthy.
3. Elaborate on the two thousand PRG code names and links. Coordinate contacting people with Les Earnest, the Stanford Alumni database and the search engines.
4. Maintain current survival status, *de mortua nil nisi bonum*. Avoid getting listed as “GONE” for as long as possible. Write your own NYT obit and send it to John Markoff (as well as cc: to bgb Baumgart at mac.com).

Chapter 4

Topics, Projects and PRJ codes.

4

On the SAIL file system each file belonged to one programmer PRG and to one project PRJ as recorded in the master file directory, the MFD. For file access, the PRJ codes were used by many people somewhat like Unix groups to partition their work into different disk areas. Outside the operating system, the accounting system tracked PRG code time sharing and disk space usage and using a database distinct from the disk file system project table assigned billing amounts to PRJ codes.

4.1 SAIL project code preview

The default login PRJ code was the numeral one, I would often login as 1,BGB. Most programmers were free to enter whatever PRJ code they liked at login

to create a new PRJ directory areas to group their files like we do today using folders. Sometimes work groups of people used the PRJ code as you would a GNU/Linux group name. For example, when a Stanford class was issued accounts for their course work, their login PRJ number was their course number.

4.2 SAIL Topic Overview

In 2007 Les Earnest wrote citing a 1973 summary -

Before getting into the sordid history of data archiving problems encountered in SAIL, let me¹ review some of the reasons for saving their records, specifically listings of research and accomplishments that we like to brag about. Here are the main research topics undertaken at SAIL in the late 1960s through '70s and some of the resulting spin-off technologies.

SAIL Research Topics [Earnest73]

- Robotics
 - Vision
 - Mechanical assembly
 - Vehicle navigation & guidance
- Heuristic programming
 - Theorem proving
 - Automatic program generation
 - Symbolic computation
 - Board games
 - DENDRAL
- Theory
 - Mathematical theory of computation
 - Representation theory
 - Grammatical inference
- Natural language
 - Speech recognition
 - Semantics
 - Machine translation
- Planetary image processing
- Computer music synthesis

The remainder of Les Earnest's Y3K paper is here, and his Y3K video talk is there.

4.3 SAIL file system project codes

4.4 SAIL accounting system project codes

4.5 Project areas with archive specialists.

In recent years the following SAILDART project areas have received some attention as computer history. Here are the active areas and the names of the

¹Les Earnest

specialists and enthusiasts. Consider this section as an invitation to contact people who have expressed an interest in a narrow SAILDART history topic that you are interested in. Be warned that the SAILDART archive is just a small splinter in the history of the big topics such as Artificial Intelligence, LISP or Robotics. The SAILDART is only of middling size within the subject of DEC 36-bit software.

4.5.1 PUB

Larry Tesler

4.5.2 TEX

Arthur Keller

4.5.3 LISP

Paul McJones

4.5.4 Music

Andrew Nelson

4.5.5 PARRY the Paranoid Simulation

Adrian Cornforth

4.5.6 Spacewar

Most recently the authors.

4.5.7 Mathematical Theory of Computation

David McQueen

4.5.8 System and WAITS

Baumgart is working on SYSTEM, Gorin and Frost are working on WAITS.

4.5.9 SUDS Stanford University Drawing System

4.5.10 Foonly

I am surprised at how many people were touched by Foonly, since in 1979 when I worked near the F1 Foonly at III it was at a low point with just five people (if you include Dave Dyer and myself). Before and after that time there were more Foonly people. The three Foonly principals are Poole, Petit and Holloway; aka



Figure 4.1: Thomas Watson at the console of the IBM-701 in 1952

Dave, Phil and Jack; aka DWP, PMP and H for Holloway. Single character PRG codes were an MIT A.I. Lab fashion for the innermost wizards such as G Greenblatt and H Holloway. The [F,*] SAILDART files that are Foonly people include PRG codes DWP, PMP, H and then AK, FW, BO, TAG, TVR.

AK is Allan Kotok, who wrote a six page, Foonly Blurb, dated 30 August 1972. The document in HTML is at [FOONLY.BLB\[F,AK\]](#) and as a scanned PDF from line printer LPT paper is here. The latter illustrates the characteristic appearance of horizontal character glyph alignment due to the drum hammer timing, study the line numbering on the left, columns one and three hammers fired earlier than column two.

4.5.11 36-bit computers

Before DEC there were IBM 36-bit computers such as the IBM-7094 from which LISP got its CAR and CDR, as well as the one of its kind TX2 at Lincoln Labs on which my mentors Les Earnest and Ivan Sutherland started their careers. Although I now have several books concerning 36-bit computers prior to the KA10; I do not think this document is the proper place for reviewing them; Except to point out the photograph of Thomas Watson,Sr. at the IBM-701 in 1952.

4.5.12 The Stanford CART

Earnest, Schmidt, Baumgart and Moravec.

4.5.13 Earnest : FINGER and FONDLE.

4.5.14 Frost : AP news.

4.5.15 Gorin : DART, DSKUSE, ACCT, RALPH, LOGIN.

4.5.16 Baumgart : GEOMED, CRE, EDFONT, XIP, XAP.

4.6 Stanford University Course Work

CS206 LISP

CS222 Music

CS220 Music

CS204 Programming

CS227 Robotics

4.7 Orphan Project Areas

For example the S1 people haven't surfaced yet to look at or comment on their SAILDART material.

S1

SUN MICRO SYSTEMS

CISCO

MARS image processing

4.8 Social Studies

Volley Ball

Spring Orgy

Prancing Pony

Sauna

Zotts

4.9 Gender, Race and Class Topics

On the edges of the SAILDART collection are remarks on the issues of gender, race and class. At the SAIL reunions it is obvious that we are almost all old white males and that we are adequately rich. Looking at the archive there is a bulletin board thread on date rape, but little or no discussion of race or economic class; except for the Les Earnest story about writing *mongrel* when applying for a security clearance.

4.10 Exercises

1. one

2. two
3. three

Chapter 5

Hardware



5

*This chapter is a triplet: **outer** system diagrams span two decades, **middle** surveys the major devices in 1974, finally **inner** is wizard craft for the chapter-17 re-enactment.*

A neat thing about old SAIL hardware is that old documentation speaks for itself. All that is needed available here are my notes on what is relevant to running 1974 software. Building a SAIL system emulator rapidly hits details which I failed to learn in the 1970s as well as mechanisms which are now irrelevant to showing off the software. For example, parity checking of the system core image, refreshing of the vector display to maintain a page of text on a screen, busy wait loops and spoon feeding work slices through the clock interrupt service are all now irrelevant to running SAIL software.

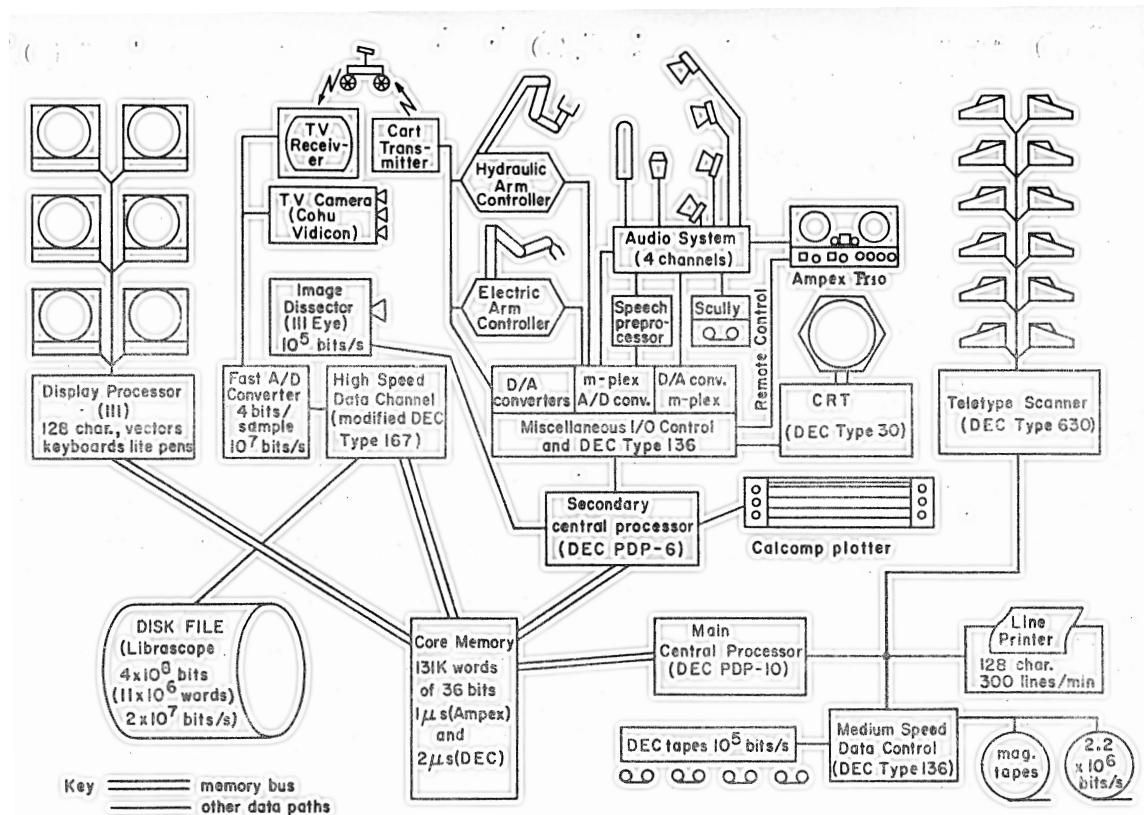


FIGURE I-1 THE STANFORD ARTIFICIAL INTELLIGENCE PROJECT COMPUTER SYSTEM

Figure 5.1: SAIL system diagram 1969.

5.1 Block Diagrams

Computers The main-frame time-shared SAIL computer system spanned three generations of DEC 36-bit machines named the PDP-6, the PDP-10 KA and the PDP-10 KL. SAIL did not use the final generation machines, the DEC PDP-10 KI, which was less powerful than a KL, but much cheaper. Special purpose processors such as the SPS-41, IMLACs, the IMP, various PDP-11s, the Samson Box, and latter day Foonlies where all peripheral to SAIL's central time sharing system.

In addition, several major computers were designed at SAIL named Super Foonly, the Livermore S-1 supercomputer, the first SUN workstation and the first Cisco network switch. The schematic drawings and documentation for the new computers and their peripherals overlap the documentation for some of

the actually implemented local hardware.

Peripheral computers existed at SAIL such as the PDP-11, a high performance FFT box, the IMLACs, the IMP and so on.

Terminals The SAILDART era spanned the transition from hard-copy terminals, teletypes with paper, to display terminals which were initially vector graphics CRTs and then video raster terminals that were refreshed from a central video disk. Burdened with an incompatible character set, SAIL was late in finally supporting the generation of 24x80 character video display boxes.

TT teletypes

The console teletype was a model 35 Teletype. I recall that the slightly fancier model 37 Teletype with lower case was available as well as a number of model 33 teletypes some of which were used as home terminals at 150 baud and later 300 baud.

Triple-I vector CRT displays

DD video raster displays

In 1973 the first Data Disc terminals were installed at SAIL providing green phosphorus screens on each person's desk.

The telnet program was implemented on the earliest versions of ARPANET and provided terminal access to SAIL for remote network users via the IMP.

IMLAC

The IMLAC display computer was used as a home terminal by Professor McCarthy and Lynn Quam,

Printers: LPT and XGP LPT:

XGP: Xerox Graphics Printer

Disks: Librascope and IBM-3330 Tapes: DEC tape, MAG tape, paper tape
Kludge Bay A to D converters

D to A converters

Petit calendar clock

Voice Synthesizer

Video and Audio Switch Robotics – Hand-Eye and Cart Eyes

Arms and Hands

Cart : Les Earnest narrative concerning [NB*link] the Stanford Cart, completely omits my humble solo participation on the Cart project between Rodney Schmidt and Hans Moravec.

Hans Moravec 1974 proposal for The Cart Project which contains the insightful paragraphs:

Cart Project 1970-1973 Baumgart

- Hans Moravec 1974 remarks

Baumgart decides he likes the idea of a robot that reasons visually, and concocts a grand scheme in which every scene viewed by the camera would be related to a model of the lab and surrounding territory. He notices the uncertainty in the analog link, and decides to make it into a digital one. This is his first digital design effort, and the result, which provides for on-off control of the motors and has no indication of the orientation of anything, is considerably inferior to the original in concept, and in addition works unreliably. The original servo electronics are disassembled or misplaced, making his changes irreversible. He rationalizes that the problems with the link are unimportant, since, when his visual reasoner works, it will be able to deduce the state of things, and detect when a transmitted command has failed, to try again.

The enormity of the effort needed to make his plan a reality becomes apparent to him as he works on sub-problems. Since it would become possible to actually use a vehicle only when his proposed scheme was almost completed, and since he now sees that it is unrealistic to think that it could be brought to fruition in a reasonable number of years, he abandons any serious efforts directly concerned with the cart, but maintains his association with it, as a status symbol and a toy. He occasionally drives it around for show, often over rough ground, contributing to its mechanical decline. During this time several other graduate students are steered towards this essentially nonexistent "cart project". They are disillusioned by the lack of a coherent plan and suffer from too little guidance and from conflicts with Baumgart's personality. All these associations are short lived and unhappy. Baumgart finds success and happiness working on the graphics and vision sub-problems suggested by his original concept.

5.2 1974 Devices

5.2.1 CPU

The main computer at SAIL in 1974 was a PDP-10 model KA. Its Central Processing Unit has two device codes PI: and APR: for controlling the Priority Interrupt and the Arithmetic Processor.

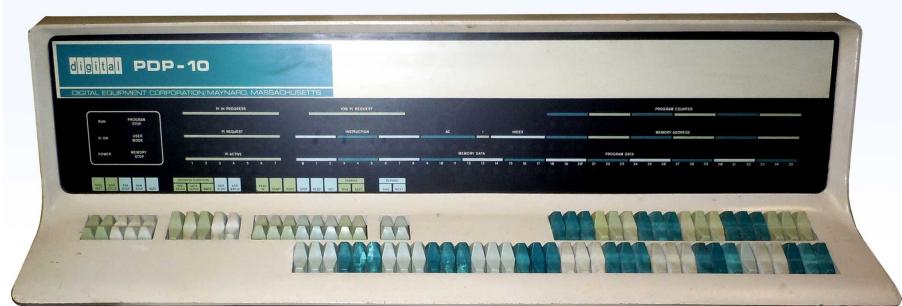


Figure 5.2: Switches and incandescent lights on the CPU console.

The plastic rocker switches were chosen by Ken Olsen (CEO of Digital Equipment Corp.) and the blinking lamps were tiny incandescent bulbs not LEDs.

5.2.2 CTY

The console terminal was a teletype at device code CTY: with a mechanical print head, black ink ribbon and a roll of paper.

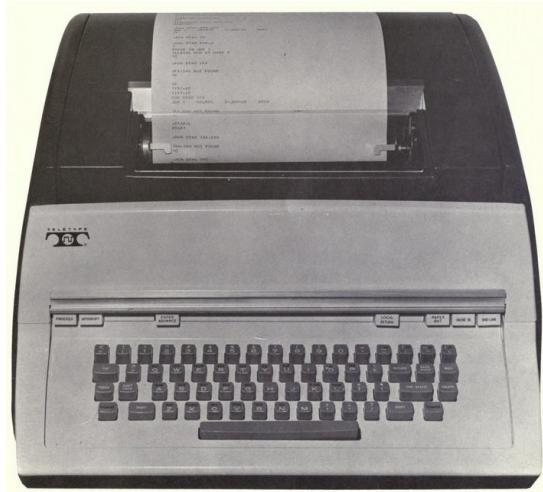


Figure 5.3: Model 37 Teletype CTY:

5.2.3 SYS

The shared online SAIL file system in late 1974 was implemented on three disk packs of a four drive IBM-3330 storage system smaller than the one in the picture. The IBM specification advertised 200MB per pack, the SAIL operating system raw capacity was 609MB using four such packs, 3 packs for SYS and 1 pack was called UDP.



Figure 5.4: SYS: on an IBM-3330

5.2.4 DKB

Display terminal keyboards SAIL custom ordered from Microswitch.



Figure 5.5: Stanford Microswitch Keyboards

5.2.5 III

Ralph Gorin at a Triple-I vector display terminal



Figure 5.6: REG at III

5.2.6 DD

A typical Data Disc raster display terminal video monitor on a professor's desk (*actually DEK's desk*).



Figure 5.7: SAIL keyboard and DataDisc video monitor

5.2.7 LPT

The shared line printer, device LPT, was manufactured by Data Products OEM to DEC. The Stanford line printer had a custom drum which spins the full font

set of 133 glyphs past a bank of 128 solenoid hammers, one hammer at each of the 128 column positions. Hammer timing is critical and early Stanford documents were noted for sloppy horizontal character alignment.

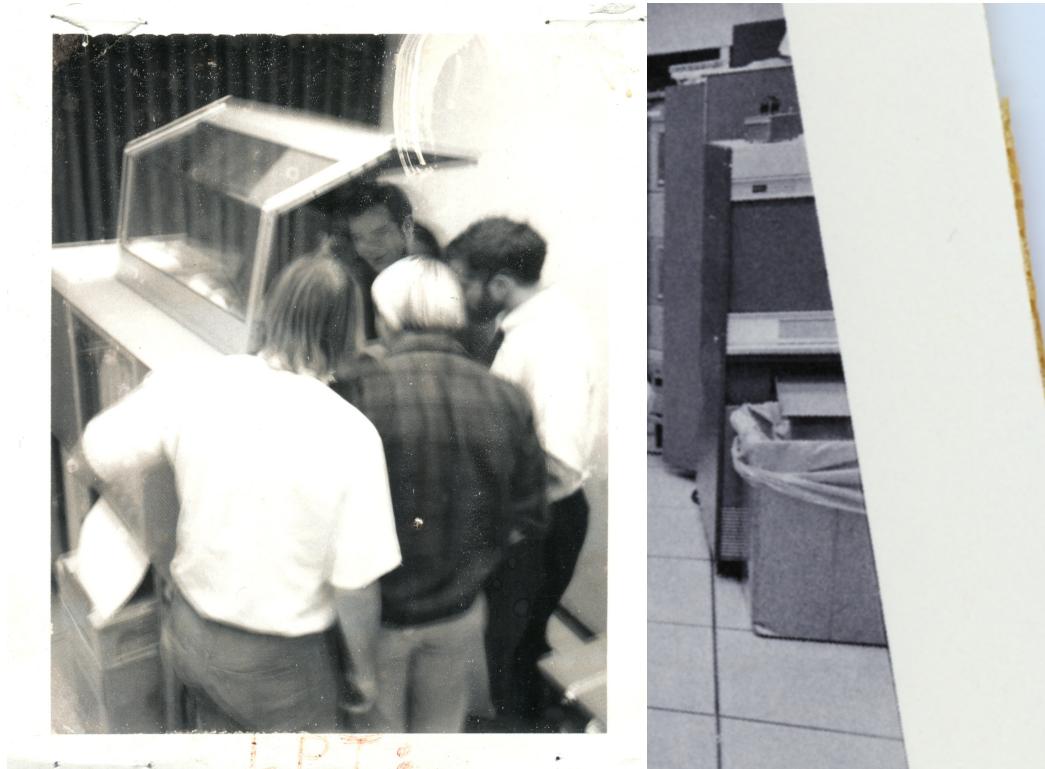


Figure 5.8: LPT and XGP

5.2.8 XGP

The XGP: Xerox Graphics Printer. This fragment is the only picture I have of the XGP in situ at SAIL in 1972. The XGP output falls directly into the wastebasket for users to retrieve. The computer room floor tiles were two foot square, so the XGP at SAIL was smaller than the XGP prototype at Xerox Webster and later at Xerox PARC.

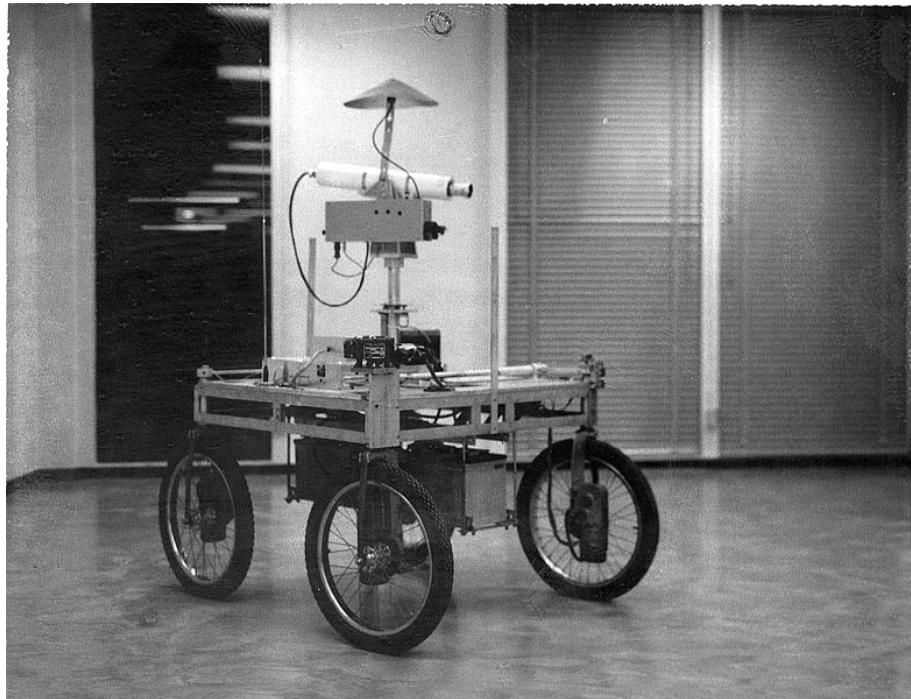


Figure 5.9: Cart 1975

5.2.9 HAND EYE

5.2.10 CART

5.2.11 KLUDGE BAY

5.2.12 LIBRASCOPE

5.2.13 MAG TAPE drive devices MTA and MTU

The two standard tape drives, as device MTA, at SAIL in 1974 were 7-Track and could write 800 BPI onto 2400 foot reels of tape. So 2400 feet times 12 inches times 800 BPI Bit-frames-Per-Inch of 7 bits (6 bits data and 1 bit parity) with the 36-bit PDP10 words taking 6 frames, there is dead space for record gaps, and so optimistically one reel of 7-Track tape holds up to 15 Megabytes.

Late in SAIL history, 9-Track tape drives were installed as device MTU. The 9-Track drives could write 6250 BPI onto 3600 foot reels; so 3600 feet times 12 inch times 6250 bit-frames-per-inch, 9 bits per frame (8 bits data and 1 bit parity) with 36-bit PDP10 words now taking 5 frames, there still are record gaps, and so optimistically one reel of 9-Track tape holds at most 180 Megabytes.

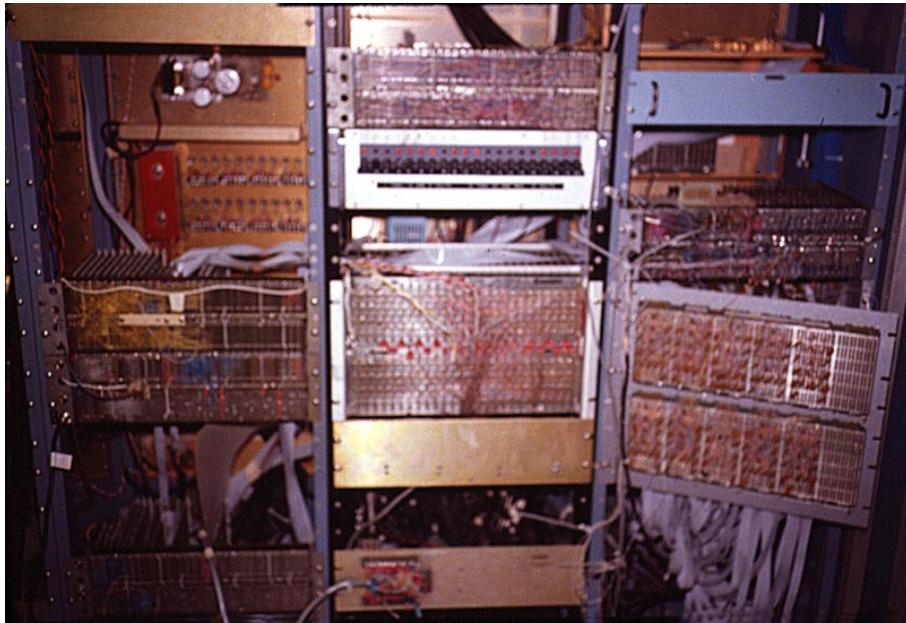


Figure 5.10: Kludge Racks 1979

Capacity	2400 feet	3600 feet
800 BPI	15 MB	22 MB
1600 BPI	30 MB	44 MB
6250 BPI	120 MB	180 MB

5.3 Wizard Hardware Lore

Do not meddle in the affairs of wizards for they are subtle and quick to anger. - Tolkein LOTR

CPU The easy part of learning the PDP-10 machine code is the seemingly simple Effective Address Calculation and the illusion that the instruction set is an orderly array of 512 opcodes in eight major groups.

CONSOLE TELETYPE foo

PETIT REAL TIME CLOCK

MICROSWITCH KEYBOARD SCANNER

SYSTEM DISK By disassembling all the versions of SYSTEM.DMP and WAITS.DMP found in [S,SYS] I can see that the symbol LSTTRK (*last track*) definition for the total file system disk capacity (*minus one track*) changed overtime as follows:

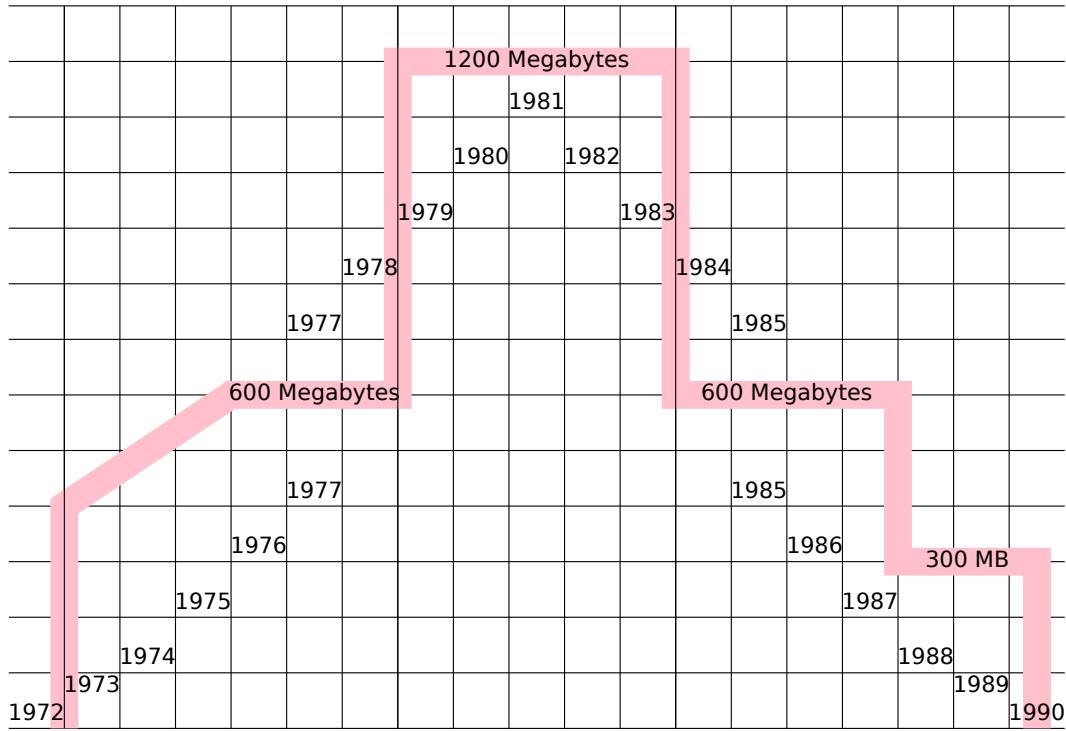


Figure 5.11: SAIL online disk CAPACITY by date

Date	[S,SYS]	version	LSTTRK	2336. PDP10 words per track	CAPACITY	days	years
1 1972-11-04	SYSTEM.DMP	~1~	0073277	octal 10512. bytes per track	304 Megabytes	94	~ 0
2 1973-02-06	SYSTEM.DMP	~7~	0131037	= 30399. tracks = 45599. bytes = 457 Megabytes	506	~ 1	
3 1974-06-28	SYSTEM.DMP	~30~	0166577	= 60799. tracks = 639119088 bytes = 609 Megabytes	424	~ 1	
4 1975-08-26	SYSTEM.DMP	~56~	0420144	= 139364. tracks = 1464994368 bytes = 1397 Megabytes	19	~ 0	
5 1975-09-14	SYSTEM.DMP	~58~	0265355	= 92909. tracks = 976659408 bytes = 931 Megabytes	38	~ 0	
6 1975-10-22	SYSTEM.DMP	~59~	0227160	= 77424. tracks = 813881088 bytes = 776 Megabytes	1119	~ 3	
7 1978-11-14	WAITS.DMP	~2~	0323552	= 108394. tracks = 1139437728 bytes = 1086 Megabytes	136	~ 0	
8 1979-03-30	WAITS.DMP	~8~	0361747	= 123879. tracks = 1302216048 bytes = 1241 Megabytes	1738	~ 5	
9 1984-01-01	WAITS.DMP	~58~	0227160	= 77424. tracks = 813881088 bytes = 776 Megabytes	1472	~ 4	
10 1988-01-12	WAITS.DMP	~114~	0036174	= 15484. tracks = 162767808 bytes = 155 Megabytes	817	~ 2	
until							
1990-04-09	WAITS.DMP	~141~					

It is difficult to feel what it was like in 1974 when disk file storage was both so very expensive and so very small. Note that the SAIL system first went into decline 1984-01-01 as its file storage space was cut.

VECTOR DISPLAY The character display was analog driven and the III glyphs corresponded to the raw (mode 100) line printer drum codes except for 011 none, 013 integral sign, 014 plus-or-minus sign, and 0177 backslash.

RASTER DISPLAY

XEROX GRAPHICS PRINTER

LINE PRINTER The six extra glyphs were center dot 000, gamma 011, small delta 012, integral sign 013, plus-or-minus sign 014, and circle-plus sign 015 each requiring a 0177 prefix. I suppose we may assume that 0177 followed by 0177 prints as one backslash.

5.4 Exercises

1. Isolate the hardware transitions that most effected user software.
2. Review and quantify reliability over the decades.
3. Chart the quantity and ratios of CPU power, cache-memory, main memory,

Chapter 6

Software

6



Software straddles *Context* and *Content*. Software as *Context* is that the Time Sharing System and its Utilities such as DART were the tools that recorded the content. The content recorded included the system software itself as well as tools for authoring, compiling, loading, testing, maintaining, repairing, training and allocating resources to run the software.

Software as *Content*, viewed from 2016, is the surviving set of old SAIL files. The old SAIL files are a heap of shattered pottery shards. The software pottery fragments need to be inspected, sorted and re-glued together to form program vessels that have enough shape to be assembled, compiled, loaded and executed in the present. There is no computer museum (as of 2014) with

PDP-10 hardware that replicates the SAIL environment sufficiently to restore DART tapes (or virtual DART tapes) to run the software necessary to reassemble source files into machine executable images. There is no emulation software (as of early 2014) extensive enough to run the SAIL binary code near its totality. I have succeeded in running my thesis program GEOMED in single user mode. I have succeeded in running a 1974 SYSTEM.DMP for the dual PDP-6 and PDP-10 through its initialization, startup greetings, a few simple console commands, and its executive DDT debug tool.

A clue to long term preservation, *or just a fluke*, some of the old LISP programs (e.g. ELIZA) run almost as-is on 2014 computer servers – once you get the LISP source text translated out of the SAIL 7-bit non-standard ASCII. But I have digressed from the software that preserved SAIL which is a program named DART running on a computer Operating System named SYSTEM which in 1976 was renamed WAITS. To revisit this cyber world requires understanding the tools that built and ran it.

6.1 Backup Utility: DART

Like in any Douglas R. Hofstadter book¹, DART is seen looking at itself. DART even looks at Hofstadter who had an account at prg code DRH. The DART program wrote the tapes that preserved the DART program as well as 18 years of DART catalogs of what was written on the tapes. As well as logs, plans and discussions concerning MCOPY version #6 of DART which was the final master copy of the three thousand low density tapes into the final high density 229 reels.

I bootstrapped myself into this *strange loop* starting with a Unix tape reading program from Marty Frost that was able to read SAIL files off DART tapes into SUN Unix workstations. However that program lacked important details concerning the SAIL-WAITS file system and the DART recording format. In 1997, I was pleased to find my own files [*,.BGB] on the earliest few tapes. Then later, in early 1998, Professor John McCarthy encouraged me to go back and read all the tapes to disk. We (Frost and I) were almost too late, since we had to rescue the tape drive (a full six foot tall rack) from a junk heap at a loading dock of the Paul Allen Building catty corner across the street from the CSD Wm. Gates Building. The junk was about to be recycled. We rolled that tape drive back across Serra Street to the basement computer room of the Gates Building. We simply read the tapes “as-is” without any conversion using the Unix utility command ‘dd’ and a small amount of bash scripting to roll the raw tape records from the ‘dd’ copy into large compressed tar tgz files, one tar ball per reel. Reading a reel would take 15 minutes. We were on our own time outside of paid work hours. The tape drive read heads were in poor condition and required frequent cleaning.

¹one famous Hofstadter book is *Godel Escher Bach*, another Hofstadter book is *I am a Strange Loop*

Starting with the Unix read tape 'C' language program from Marty and making some changes based on hard-copy manuals (the PDP-10 reference manual, the FAIL assembly language manual, the system API UUO manual and the system command Monitor manual) which I had kept paper copies from the 1970s, I was able to read the source files of the PDP-10 assembly language DART program and its documentation which then made it easier to write improved conversion programs with better consistency checking and error recovery. That sequence of programs are name 'undart-something.c' for example a recent one is named 'undart-2014.c' which converts the raw DART tape records into the data blobs with metadata database CSV file indicating what SAIL file-names, SAIL programmers, SAIL projects and dates are associated with each data blob.

DART was first written by Ralph Gorin, version 1972 to 1979, then taken over by Marty Frost.

The saga of 'undart-2014' is recited in chapter #9 when we *exegesize* the DART meme for the benefit of digital archive specialists.

6.2 Operating System: SYSTEM and WAITS

Before Windows, Apples and Unix; before men went to the moon in 1969; very few computers had time sharing systems. The time sharing system at the Stanford A. I. Lab was called SYSTEM and it had been derived from a Digital Equipment Corporation system which was called MONITOR. In 1976 after a naming debate, the SAIL operating system, née SYSTEM was renamed WAITS. As mentioned in chapter one, the WAITS acronym was never officially nailed down. I propose "West coast A. I. Time Sharing" will do as a mnemonic, in contrast to the East coast sister system named ITS (Incompatible Time Sharing) at the MIT A. I. Lab. The hubris of incompatibility proved fatal to each.

The SAIL system "shell" and its in-line command text editor.

The SAIL operating system command "shell" was entangled (buried in clock service CLKSER) inside the operating system. The addictive text input in-line editor named "LINE" was wired into the keyboard device driver of the operating system. Latter day LINE appears today in GNU/Linux bash as the GNU **readline** library and the low level character edit commands in emacs or vi. Larry Tesler first at SAIL but then later at Xerox PARC coined the terms "cut" and "paste", which in TECO and EMACS were called "KILL" and "YANK".

6.3 File system repair: RALPH

The eponymous program RALPH written by Ralph Gorin is the SAIL-WAITS disk file system checker. Reading the source code of RALPH.FAI[S,SYS] is

how I deduced the block level disk format for the SYSTEM disk image of 1974. The best latter day RALPH source is at RALPH.FAI[S,SY] dated August 1983 and the final version in 1974 is at RALPH[ACT,REG]21. Reading the SYSTEM source code itself for DSKSER and DSKINT is at a yet lower level and the file system format which is buried in the UUO service for routines that implement OPEN, LOOKUP and the I/O transfers. RALPH uses a back door called GOD mode to read and write directly to the disk.

6.4 Help

A user level program that can be run without being logged in is HELP. The HELP source code leads to the documentation on how to use the 1974 system.

6.5 Login and Logout

Nothing interesting can be done until you get past the gatekeeper, LOGIN.

6.6 Accounting ACCT *SPY*

Another Ralph Gorin program which existed and ran in conjunction with the Operating System was named ACCT and was visible to everyone as the JOB-NAM “*SPY*” which was always there.

6.7 Assembly language: MACRO, FAIL and MIDAS

MACRO was the early DEC assembler. FAIL, an acronym for Fast A. I. Language, was the PDP-10 machine code assembler at SAIL. The rationale for FAIL was both for performance efficiency and the ability to add and control additional features. On latter day DEC PDP-10 systems there was an assembler named MIDAS which was relevant at SAIL for ARPA network software and for people who learned PDP-10 MIDAS before arriving at SAIL.

Further assembly language tools include LOADER, DDT, RAID, and CREF.

6.8 The LOADER

The loader creates executable DMP files from REL files. Strange that the early DEC loader continued at SAIL almost unmolested, although Tovar (TVR) did replace the symbol table lookup with a binary search rather than its original linear one.

6.9 The Text Editors 'E'.

During this period, the tools for authoring text for software as well as for documentation transitioned from the neolithic primitive (EDIT and TECO),

through the mesolithic usable (first Stopgap, then Son of Stopgap SOS and TVED) until 'E' which dominated the SAIL epoch. Later emacs arrives in the late 1970s. The point here is that the text editor named E and the E format dominate the period that is central to the SAILDART archive. An E document always had a block index table of content at the front of the file.

6.10 RPG - Rapid Program Generation

The “IDE” concept at SAIL was RPG to invoke edit, compile, execute and debug.

The SAIL work flow was not what is now called an IDE, Integrated Development Environment. However in forgiving retrovision the primitive SAIL IDE consisted of RPG, the Rapid Program Generation suite of commands.

6.11 Software content payload preview

The **high level programming languages** at SAIL were

- LISP
- MLISP
- UPLANNER
- SAIL
- PASCAL
- 'C'
- BASIC
- F4
- FORTRAN
- ADA

Some of the **FONT software** programs were

The dominant **document formating markup languages** were PUB and \TEX. The earliest SAIL document formatting programs were XIP, XAP, XGP and POX.

Graphical Design Software,

- SUDS suite
- GEOMED for 3-D models

Games and Puzzles

- Spacewar
- Life
- Adventure
- Go
- Kalah

Robotics

Image Processing
The Hand-Eye Robotics Library

Music, Audio and Voice

John Chowning, Leland Smith **SCORE**. Andy Moorer. Jim Grey.

Exercises

1. Experiment: Find out how hard is it for a non-SAIL person to become a code reader of SAIL system PDP-10 assembly language.
2. two
3. three

Part II

Content of SAILDART²

²The banner means “*A Discussion about the Content of the SAILDART*” and where you can find the actual content.

Part II Content of SAILDART

PART



chapter	7	Access
8	Provenance	
9	Exegesis	
10	Taxonomy	
11	Corpora	
12	non-DART	

The next four chapters, 7 to 10, are archive technical. There is one visual moment that justifies two photographs to document moving the magnetic tapes from Gates to Green. Chapters 11 and 12 will be heavily illustrated for a coffee table edition, arrays of thumbnails will suffice for this white paper edition.

From Data to Information.

Primary content of the SAILDART web site is the set of SAIL computer files from the DART backup tapes.

Secondary content is the DART metadata about the SAIL file system which appears both out-of-band in the DART tape records and in-band inside the accounting files that DART generated in the course of performing its backup runs, which were then backed up in turn by later backups. Strange to relate, the final months of the SAIL Time Sharing System were spent just running the DART tape conversion.

Tertiary data on the SAILDART web site is other data closely related to SAIL that is not found on the DART tapes. The tertiary data includes scanned documents, video from reunions, as well as post DART writings such as this white paper.

Chapter 7

Access, Privacy & Search

7

There are two implementations for working with the SAILDART archive. The ***second*** is the **public** SAILDART web site, which is built from the ***first***, a **private** GNU/Linux file system.

7.1 Public Access

7.1.1 Access by Canonical URL

The canonical and permanent, SAILDART file URL is simply the old SAIL PDP-10 file name, extension, project, programmer with the old punctuation marks optionally postfixed with a decimal revision number (*there are no curly braces around specific revision numbers*).

```
FILNAM[PRJ,PRG]
or
FILNAM.EXT[PRJ,PRG]
or
FILNAM.EXT[PRJ,PRG]{revision number}
```

To get a bitwise exact copy, append "_octal" to the URL. For example:

```
wget -q http://www.saildарт.org/BUCK75.FNT[XGP,SYS]_octal
```

serial numbering the data blob hash codes.

7.1.2 Access from programmer Home Page

For each PRG code (well actually PRG+1 owner codes, which are equal to PRG codes for most everyone except when a code was reused for a different person) a triple frame SAILDART home page exists at URL

```
http://www.saildарт.org/BGB
or
http://www.saildарт.org/[1,BGB]
```

7.1.3 Access by Date

I once had the SAILDART files accessible by URLs in the form of

```
www.saildарт.org://{isodate}/FILNAM.EXT[PRJ,PRG]
```

for accessing a revision without having to know its {revision number} since server side mechanism could select the correct revision existing on the given date. This is not a unique canonical URL, but rather provides a large set of URLs for each day in the span of the file revision's existence. I could be encouraged to re-implement this form of access, and have appended it as a low priority exercise.

7.1.4 Access by Serial Number of content blob

7.1.5 Access by pathname

7.2 Copyright and Ownership

The copyright status of the almost one million items inside the SAILDART archive varies and may be looked up per item. Most SAILDART items were never published, others are public domain. SAILDART is an archival collection with human curators. Compliance with the original ARPA, NSF and other contracts supporting academic research at Stanford University is continued best effort. Compliance with the Stanford University policy for archiving research data continues.

7.3 Privacy, Courtesy and Ownership

John McCarthy punted on the privacy issue. He said to me, *paraphrasing* 1. Do not be in a hurry to contact Stanford officials, 2. Get advise from Les Earnest and Marty Frost, and memorably he repeated the cliché: 3. *It is easier to ask for forgiveness than it is to ask for permission.*

Stanford University has had continuous possession of the DART permanent tapes. The 229 reels of DART tape are now safely housed in the Digital Collection at the Green Library, by my initiative with the assistance of Earnest, Frost and Hartwig.

7.3.1 Who guards the guardians?

Quis custodiet ipsos custodes?

Peer pressure. The guardians must guard each other. All of us must take care of everyone. Les Earnest has observed that at any computer project there is an inner circle of system programmers who have access to everything.

7.3.2 Stanford Research Policy Handbook

This URL¹ links to a page concerning Stanford University policy on the retention of and access to research data. I am aware of this policy now, and I was aware of the issues and ambiguities of an unsorted bulk data collection in 1998 when working with John McCarthy and Ted Selker on long term digital preservation for data mining at the IBM Almaden Research Center. From the policy, I wish to quote four sentences verbatim:

- #1. When individuals involved in research projects at Stanford leave the University, they may take copies of research data for projects on which they have worked.

¹<https://doresearch.stanford.edu/policies/research-policy-handbook/conduct-research/retention-and-access-research-data>

#2. Original data, however, must be retained at Stanford by the Principal Investigator.

#3. Research data must be archived for a minimum of three years after the final project close-out, with original data retained wherever possible.

#4. Beyond the period of retention specified here, the destruction of the research record is at the discretion of the PI and his or her department or laboratory.

I will claim a wide interpretation for sentence #1, starting with my PhD thesis work on which I indeed hold a 1974 copyright and which arguably is my intellectual property and not that of Stanford University. I am in compliance with policy sentence #2 since the original media is still at Stanford. John McCarthy seemed aware of the Stanford University policy ideas in sentences #3 and #4, and he took it that some folks might exist that assumed the three year retention period was a maximum after which old data should be destroyed in order to avoid difficulties and to cut off the possibility of belated reviews or whistle-blowing. John McCarthy was of the opinion that AI should be like Astronomy where research records are kept forever.

7.4 Search

7.4.1 External Search Engines

The SAILDART collection that has been on the web for the past decade is too large, too fragmented and too redundant for the search engines to make much sense of it. The search engines downgrade sites that are as large and as illegible as SAILDART has been. However search for keyword SAILDART appended with a couple of your special keywords will turn up SAILDART stuff. For example, search “SAILDART ZORK” returns a set of SAIL files referring to Don Woods game Adventure.

7.4.2 Internal search mechanisms

The digital curators who have a copy of the SAILDART in a file system can navigate the million files readily using find and grep. I have built and used a full word index (a concordance) from time to time, but I do not have one built at the moment. Frequency histograms of N-glyphs and N-grams are a routine way of finding stuff that looks like given text, however I do not have a SAILDART search tool kit to hand off yet. Semantic networks of documents using the same vocabulary (especially surnames) would be next.

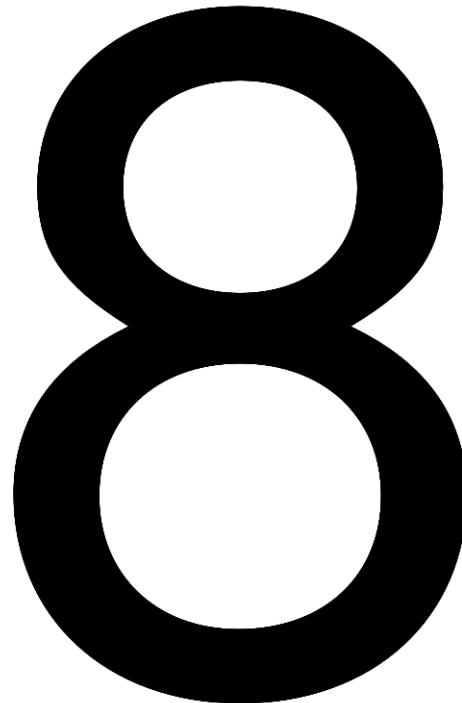
7.5 Digital Curators

7.6 Exercises

1. Build (rebuild) a new suite of concordance tables by Words, N-grams, Names, phrases, sentences. Letter (character) frequency 1s, pairs, triples, N-characters.
2. Finish writing taxonomy predicates, such as `isLISP` and `isFAIL` using either a parser or frequency histograms or both.
3. Rebuild the *search by date* mechanism into the SAILDART web presentation.

Chapter 8

Provenance



As I see it, provenance of a digital archive has two parts, first is the recitation of the chain of custody of the media, and second is the fidelity of the data transcription into working copies for preservation, circulation and presentation. Provenance becomes complex when the original media has been lost or intentionally replaced and we must iterate on media custody and data fidelity for each copy event. Unlike Euclid's Elements or the Bible, this archive is not analog and has not yet needed to be translated by human scribes into Arabic

or Greek, and unlike those examples this information was born digital, copied only twice by 20th century electro-mechanical means, and is yet again digital. Also noteworthy, over the past 40 years, the 50 Gigabyte quantity of the SAIL-DART has transitioned from a room sized “off-line” “big-data” set of 3000 reels of twelve inch tape (weighing 2.2 pounds each), into chip size, on which all of the data can now fit “on-line” inside one CPU main memory address space. The quality of the two off-line copy events of 1990 and 1998 will be depicted in this chapter. Future digital copy events will be cheap, fast, frequent and bitwise exact.

After the provenance paragraphs, this chapter concludes with a description of motives and practices which prejudice which parts of the archive are visible today and who can see them in the 21st century under the existing copyright, practical triage and social politeness constraints. Past the year 2100, it is my wish that this small quantity of data shall be free and open to all. My mentors and teachers: John McCarthy, Les Earnest and Don Knuth have requested; my associates Diffie, Frost, Petit, Gorin have not complained too much; there has been very little push back from third parties; as well as very little encouragement or push forward. All has been quiet on the SAILDART web sites except for the relentless crawling of the many search engine robots, Robots-dot-Text non-compliant download attempts. And a few times each year, I receive a relevant query from a human concerning what can be learned from the SAILDART archive.

8.1 Custody

For the 229 reels of DART tape the provenance story told here will detail the path from off the 1970s SAIL-WAITS File System through the lab relocation and a tape media conversion until 1998 when the final tapes were read into 229 Unix file system compressed tar balls, tgz, each with its MD5 hash value. The 229 tar balls expand into exactly 41620 DART records -26 = 41594 DART records, which in turn “undart” into 886476 unique data blobs. Each data blob has its MD5 hash value. The MD5 hash values were serial numbered sn/000001 to sn/886476. A traditional archivist might wish to call this serial numbering the SAILDART accession numbers.

Data blobs often have an obvious MIME/type such as human authored text, using the text editor named “E”, computer generated text, digital images (usually black and white at six bits per pixel), vector graphics, executable PDP-10 machine code, audio data (often as twelve bit samples), accounting system database records and DART program backup database records. Along with the data blobs the *undart* processing generates SAIL file system metadata such as the filename, extension, project, programmer, protection bits, size, an xor checksum and four date-time stamps.

Recapping, there were twenty five years of SAIL 36-bit computer operations from 1966 to 1991; within which there were eighteen years of low density DART tape recordings, 1972 to 1990 which are serial numbered 1 to 2984. From 1988

to 1990 First baton pass: 7 to 9 track tape at Margaret Jacks. Marty Frost copied the almost three thousand reels of low density seven track tape into the 229 reels of higher density nine track tape, serial numbered 3000 to 3228. March 1998 Second baton pass: Tape to Disk at Gates. Bruce Baumgart (with the help of Marty Frost, Les Earnest, John Nagle and Tom Costello copied the 229 reels of the DART 3000 series (via external 9 GB disks) into various systems and media at IBM Almaden and at the Baumgart residence. April 2011 Third baton pass: original DART media is moved from Gates to Green.

On 26 April 2011 we (Baumgart, Frost, Earnest, and Hartwig) moved the 229 reels of DART tape from the Computer Science Department at Gates Hall to the special collections at Green Library on the Stanford Campus.



Daniel Hartwig and a box of tapes



Marty Frost and Les Earnest

Bruce BAUMGART – Statement of the custody of the physical reels of DART tape.

Based on DART tape header dates, I assume that the low density tapes reel#1 to reel#1583 were written in the computer room at SAIL in the D.C.Power Building at 1600 Arastradero Road Palo Alto CA and that those tapes were moved to MJH in November 1979. Tape reel#1584 to reel#2984 were written in MJH. The tape conversion software was developed and tested in early 1988, but not vigorously used until May 1990. Only the first three high density tapes were written in 1988, the remain 226 reels were written in 1990, apparently there was no tape conversion work done in 1989. The 229 high density tapes were moved from MJH to WmGB in December 1995 or January 1996. We read the 229 high density reels of tape using Sun Microsystems equipment to 9 GByte SCSI disks (Maxstor) that I happened to own at the time. The 9 GByte disks were Sneaker Net (that is hand carried by automobile) to the IBM Almaden Research Center where I was working as a Research Associate. The tar files off the 9 GByte disks were transferred to various systems I had access to at the time (AIX and Redhat Linux) as well as DLT tape and the ADSM backup system. I still have a 1998 set of gold colored CD disks with the 229 tar files.

Reading one reel of tape took 15 minutes and would leave a noticeable

quantity of iron oxide dust on the tape read heads and in the tape path so we would clean the tape drive with alcohol swabs frequently. I trust that the next readers of these tapes will have exquisite technology that avoids inflicting as much tape damage as we inflicted. We fetched and returned the tapes from a storage room adjacent to the locked server room in the basement of Gates.

While I was at IBM, the media included DLT-IV tape cartridges. Only a single DLT cartridge was needed to hold the archive at DLT model 7000 density. The SAILDART data fits on some forty (40) ordinary CD compact disks. Such sets of CD disks are slower and less convenient to read and write in bulk, but the CD readers were ubiquitous, the media is cheaper than DLT tape, and so as a long term archival strategy, writing to CDs was briefly considered a viable approach. What soon proved more viable was a chain of many cheap disk drives SCSI to IDE to SATA. The SAILDART preservation copy of the DART tapes now fits on USB thumbnail drives as well as SD memory chips. Writing a copy to the IBM ADSM (later the product was re-branded Tivoli something) proved to me the lack of endurance of large data sets in the corporate research environment. In the late 1990s at Almaden the bandwidth and backup time windows were such that only with great patience and skill could 50 gigabytes of stuff be written into ADSM and that without senior management priority such large quantities could never be read out and my large presence inside the robotic tape machine was well known and resented. I was unfortunately asked to perform a similar large backup stunt again for some of my peers in the Web Fountain group at IBM, and so I ended up building a skunk works cluster of cheap commodity disks outside the ADSM service.

PC copies to special people: At my own expense, I built three Redhat Linux PC systems with a full copy of the SAILDART, and gave them away to Marty Frost, John McCarthy and Les Earnest. Usually I avoid cute hostnames, but in this case those Redhat systems were named after American Civil War generals: Grant, Lee and Sherman; so when a Les Earnest email to me says U.S.Grant lost his whatever or failed to do something, then you will know what that refers to.

CD distribution of individual PRG areas to the authoring individuals occurred from late 1998 to 2000.

8.2 Fidelity

The bytes found on each high density tape in the 1998 reading using the GNU/Linux 'dd' utility were aggregated into 229 compressed tar balls and MD5 hashed. The hash numbers assure that the present 229 tar balls are the same as the 1998 ones.

In 2015, the GNU/Linux 'tar' dependency was removed and the raw DART byte string written into a single file.

8.3 Prejudice

The files now visible on www.saildart.org are files which were visible during the SAIL years, 1972 to 1990. Plus files from disk areas of people who have granted permission to display their files.

My first work on converting files to modern formats concerned my own files GEOMED and my PhD thesis work. Resulting in good presentation of the PLT and VID files. My recent interest has been the operating system PDP-10 assembly code software which I have narrowed down to just what is found for 1974. This is a tactic to get some results out in a finite amount of time with little or no help.

Meeting with Les Earnest from time to time, we have further decided that all DART index filenames and dates can be made public.

8.4 Infidelity

Running the 1974 SAIL operating system as an exact emulation would seem to us 21st century people as very slow and ugly, it would also crash a lot. The SAILDART code re-enactment has taken considerable artistic liberty to remove slow and ugly as well as to mitigate system crash defects.

8.5 Exercises

1. Keep an eye on Green. Visit Stanford's Green Library to verify that the SAILDART material still exists.

2. Someday re-read the magnetic tapes.

Consider the trade-off between re-reading the tapes sooner with existing (or even worse museum grade) technology or later with more advanced technology but more decayed magnetic tapes. There is no need to chisel bones out of the La Brea Tar Pit when you have X-ray tomography. Vinyl records are now read with a Laser not a mechanical needle.

Consider reading all the old non DART tapes that can be found around Stanford University and else where. Martin Frost and I only grabbed the 229 tapes we knew were the final permanent ones, the rooms and attic storage areas where these tapes dwelled in the 1990s had many other reels of tape.

Date **Monday 3 February 2116**, the 229 reels of magnetic tape are moved from the Stanford Green Library back to the hills, this time up Sand Hill Road, not Page Mill Road, to the newly opened Stanford Linear Archival Conservatory. A special building for a collection of linear media from the 20th century including celluloid films and magnetic tapes stored in stacks of lamina of carbon fiber re-enforced foam. The fragile films 8mm, Super8, 35 millimeter, 70 millimeter and magnetic tapes audio, video tapes and numerous forms of early computer magnetic media are unreeled once only unto a plank of the pellucid clear, exquisitely thin, rigid carbon fiber foam. The planks stack in the two mile long archival vault built on the site of a mid twentieth century physics project.

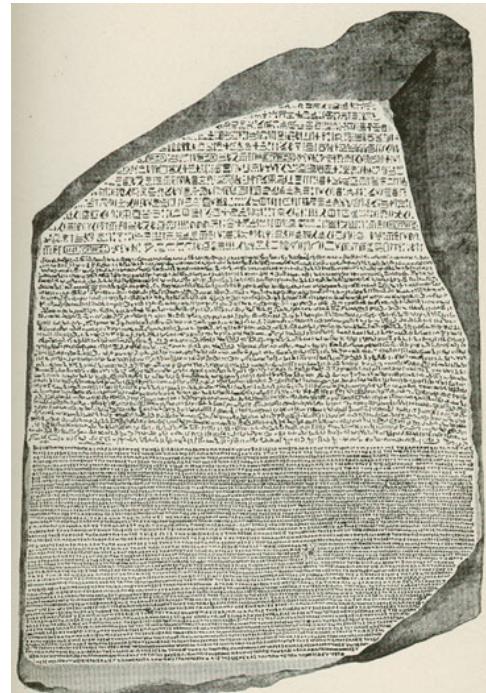
The SAILDART tapes rest shiny side down, iron oxide side up, so that the electro-magnetic scanners view the top side looking down and the optical sensor the bottom shiny side looking up. Opto-Chem is secondary for digital magnetic tape, but finds numerous greasy human fingerprints mostly near the ends of reels but on occasion in the middle of a tape along with obvious mechanical damage to the media. The fingerprint images can be dated to either the 1990 write phase or the 1998 read event.

3. Keep an eye on SAILDART. This document is a pointer to a long message, verify that you have access file (byte vector)

Chapter 9

Data Remix - Exegesis

9



I borrow the word *exegesis* from scholars who study ancient texts like the Dead Sea Scrolls. For a digital archive, *exegesis* is how to convert old formats into current ones. One gold nugget from the Power-point School of Information Technology is the slide showing a pyramid labeled from bottom to top: Data, Information, Knowledge, and Wisdom. Climbing **up** the pyramid from

fifty-three gigabytes of tape *data* reveals the *information* to found at the first Stanford Artificial Intelligence Laboratory.

The raw data itself.

Data is carried on media. The story of the original DART *media* is its provenance as recited in Chapter-8. DART data copied from the original DART media can be compressed and placed into a single GNU/Linux file. The raw data and software to unwrap it into current formats fits within sixteen gigabytes, which easily fits on a digital camera SD memory card. SD stands for Secure Digital, or *Sail Dart* !

Attached here, **you may have in your hands**, copies of the SAILDART archive along with the current GNU/Linux software to *exegesize* it. The chips are labeled on the outside in black letters on white plastic sticky tape “SAIL DART 2015”. The MD5 hash values before and after compression are:

```
a33c8fff234f0e106e2e99e219f3b7b45 dart_records_from_229_reels.tar
ecc815253d76dfffc615e311bfbb4c090 dart_records_from_229_reels.tar.lzma
```

Next, I have eliminated the TAR dependency. Time capsule will be a compress Flat DART byte stream. After that, I have eliminated the DART dependency. The time capsule or hand-off baton need only have SAIL content. DART is a message carrier, it is NOT the message. The DART story is in the Provenance Chapter as a step in the chain of historical data custody.

9.1 Narrative Description of the DART format.

Five bytes to the PDP-10 word

When writing to 9-track magnetic tape PDP-10 words of 36-bits were transferred by the I/O hardware into bytes in big endian order with the final four bits in the high order of the fifth byte.

The thirty-six bit SAIL computer words were written to tape, big endian, in five octet bytes. The fifth byte of each word has four low order data bits and four bits of high order zero-bit padding. When the tape drives loses sync, octets can be lost or inserted, and the remaining words would be garbled, however by scanning in the serial byte stream for DART record landmarks, the position of the misaligned octets can be adjusted and the remaining words of the record properly aligned.

Seven bit SAIL text encoding

Most text at SAIL was encoded in a 7-bit ASCII where the thirty-one characters after zero 000 ASCII NUL were mapped into non-Standard non-ASCII glyphs. For example 001 was ↓ down arrow. Unless you are listening to this on ear buds while flying westward across the Pacific, *Look* at the SAIL-to-Unicode table and *Skip* reading the next paragraph.

Continuing to strain this narration, I wish to recite the SAIL octal codes for the **arrow** glyphs as 001 down arrow, 027 double arrow horizontal, 031 right arrow, 136 up arrow and 137 left arrow; and which are *exegesized* into the Unicode hexadecimal codes U2193 , U2194, U2192, U2191 and U2190 respectively. The **Greek letter** glyphs at SAIL were octal code 002 for α alpha, 003 for β beta, 006 for ϵ epsilon, octal 010 for λ lambda, and 007 for π pi. The five Greek letters become Unicode u03b1, u03b2, u03b5, u03bb and u03c0. For **Logic and Math** the SAIL codes 004, 037, 024, 025, 005, 026, 016 and 017 represent glyphs for boolean AND, boolean OR, for each, there exists, boolean NOT, XOR as a circle X, infinity as the lazy eight symbol and the partial differential operator. These codes respectively become Unicode u2227, u2228, u2200, u2203, u00AC, u2297, u221E and u2202. Then a few extra mathematical **relation and horseshoe** symbols encoded at SAIL as octal 033, 034, 035, 036 and 020, 021, 022, 023 which are again in order are $\neq \leq \geq \equiv$ and $\subset \supset \cap \cup$ that become Unicode u2260, u2264, u2265, u2261 and for the horse shoes u2282, u2283, u2229 and u222A.

The point here is that even if this Prolegomenon and its Exegesis are lost (or are not provided to an archive decoding test candidate) the DART gram is not hard to interpret after the idea of 7-bit characters that are nearly ASCII is re-discovered. Then the DART record metadata is merely irradiating hiccups in a stream of text. Reading all that text provides the future Cyber-Sapien archivist (*long after the self destructive Homo-Stupids have disappeared*), with the actual software which wrote the DART message, late in its analysis the decoding expert will “see” the glyph shapes in the binary font files or in the T_EX metafont files.

SIXBIT file name encoding

The SAIL-WAITS file system is primitive, it was a tool of pioneers working on the frontier. Filenames were one to six characters, optionally followed by dot and a one to three character extension. All filename alphabetic characters where uppercase. The character codes for A to Z were six bits wide as octal 041 to octal 074. The digits are octal 020 to octal 031. The blank is zero. Filename characters on the DART media (as well as internal to the Operating System) could have any of the 64 character values.

SAIL-WAITS file system

Each file belonged to a directory specified by left square bracket project code comma programmer code right square bracket. The project and the programmer codes where each one to three characters long.

Alien to the SAIL file system is the now familiar file system concept of having content blobs separate from directory entries. On GNU/Linux file systems, one or many file path names may be hard linked to one content blob, which was impossible in the SAIL-WAITS file system. At SAIL the early disk hardware was unreliable so that a seek command was not be fully trusted to

get to the proper cylinder, head and sector of a disk drive. So the file name (directory entry), which SAIL called the Retrieval Information Block (or RIB), included the file name and was written into each data block (called a Track) that was needed to hold the file's data. And so too on the DART tapes within the sequence of FILE tape data blocks each tape block has a full copy for the Retrieval Information.

Segmentation into DART records

There are two original DART record types: Tape-Marker (Head or Tail) File-Data (Start or Continue). I have added a third record type named Gap, to passover the 61 segments of bytes which failed to decode as DART records. Previewing the data shows that all the Tape-Marker records are exactly sixty bytes long and each contains the tape reel number and a date-time stamp. The first word of each DART record has its record size. The record lengths segment the whole DART byte stream with only 63 defects, continuing after a defect requires scanning for the next sane record.

With the extreme precision that is available to latter day archival software, the DART segmentation goes as follows: the long byte vector, of exactly 56_446_334_821 bytes, contains exactly 2_937_291 short segments of which 5_486 are head-tail records, 1_886_472 are file-start records, 1_045_270 are file-continue records plus the 63 gaps. Or more colloquially, the fifty-six gigabytes of tape data have nearly three million short records which contain the data and the names of about one million old SAIL files.

Three further mechanisms need to be previewed here. First, it was the intentional DART backup policy to write two copies of each SAIL file that was deemed of permanent value to two different permanent backup reels of tape. A file found by the utility programs named DSKUSE and DART resident on the SAIL community commons SYS: disk system would be marked as archived once, then marked as archived for a second time, and then there after omitted from further archiving. So each SAIL file should appear in the dart record in two places in the tape records with the same identical content, name and date-time stamp. Second, a unique SAIL-WAITS filename will appear again (with yet two further copies each time) for each newer date-time stamped revision. Generally human edited files do not change very much between revisions. Third, it was the unintentional result of unreliable disk seeking mechanism that meant that file retrieval information including the file name was stored multiple times within the file "blocks" on the disk media. That meant that the SAIL-WAITS file system would contain multiple copies of exactly the same content of a file when a file was copied from one user directory into another. Other kinds of short files (the professional digital archivist term for these files is "turd") are generated by common utility programs in many user directories with content of no value to the historical record aside from traffic analysis. The result is that the population of 1_886_472 SAIL files in the DART halves to fewer than 900_000 different content blobs, each content blob has one to many hundreds (and for a couple of blobs even thousands) of directory entry name

tags (aka retrieval information) rows in the database table of the SAIL-WAITS file names.

FILE-START and FILE-CONTINUE tape records File-Start and File-Continue records are identical in format and in content of their file metadata. The File-Start is marked type -3 in the left half of word 0. and the constant sixbit/*FILE*/ in word 19. The File-Continue record is marked type 0 in the left half of word 0 and sixbit/*CONT*/ in word 19. So describing them both as FILE blocks they have 36. words of prefix, then up to 10240. words of data payload, then a 23. word postfix which is most often completely zero except when a few bits are tinked pursuant to observations of error conditions in the reading of the low density tapes.

The FILE metadata is sixbit/FILNAM/ sixbit/EXT/ sixbit/PRJPRG/ the length of the file in words, SAIL-WAITS protection bits, mode that the file was written, and a date-time stamp.

The file data block records seen on the high density tapes are surprisingly fat considering the computer poverty of the prior 18 year period. The explanation is that the DART data format version #3 was a final revision done to handle the massive MCOPY of the 3000 old tapes into the newer higher density ones, the format was over ambitious and had allocated many bytes of space that were never used.

HEAD and TAIL tape records All the tape HEAD and TAIL records are exactly 60. bytes long. Each contains 12. PDP-10 words. Seven of the twelve words have a fixed constant value, making the HEAD-TAIL records easy to find in a byte string, the other five words carry a date-time stamp, a checksum for 10. words of the HEAD-TAIL record, the tape reel number and the tape position in feet from the tape load point which is irrelevant to the SAIL-WAIT file system but it is amusing to know where the low density tape reel images fall within the high density tapes.

There are 41_594 tape records from the higher density tapes, which each in turn contain 1 to 100 or so small records from the lower density tapes. In total there are 2_934_700 of the small records plus the 63 gaps.

The 229 reels of high density DART tape are labeled P3000 to P3229, as mentioned earlier, the reels still exist and are kept in the Stanford University Digital Archive housed in the Green Library building on the campus in Palo Alto, California. Each tape contains high density (6250 bpi) records. Each high density record is a concatenation of records from the lower density (800 bpi) tapes which were label P1 to P2984. The letter 'P' indicated Permanent backup tape as oppose to the incremental ones which were marked 'T' for Temporary. The final reel of Permanent Tape was written 16 August 1990 and that reel of tape was copied to disk in March 1998, however the earliest file I have from that reel is time stamped 17 June 1998. The rescue of the high density tapes to disk was not a well documented process, the quantity of old tape in the basement of the CSD building was overwhelming, the speed of the tape drive

was slow, the working hours were 2nd and 3rd shift, the disks drives were nine Gigabytes each and were taken off site to copy into several other systems since there was not enough disk space available to us on a single system.

The low density reels were written over a period of nearly 18 years. The HEAD of tape #P1 is time stamped 1972-11-05T11:59 and its TAIL is marked 1972-11-05T12:23 which implies that first tape took 24 minutes to be written on a quiet Sunday in November around lunch time. Richard Nixon wins re-election to a second term as president of the United States on the following Tuesday 7 November 1972.

The high density reels were written over a period of nearly 31 months. The HEAD of the first high-density DART tape #P3000 is time stamped 1988-02-01T17:17 The TAIL record on the final high density tape #P3228 is dated 1990-08-16T22:55 so at nearly 11 PM on Thursday in mid August the DART record ends. Iraq had annexed Kuwait during the first week of August 1990.

The final lower density tape #P2984 is time stamped 1990-08-17T16:43 which overlaps the time period in which the final high density tape is written.

GAPS

The data found in the 63 gaps, is assigned its MD5 blob serial number and tagged with a unique SAIL file label and included in the SAILDART collection as allowed by KISS design authority (Keep It Simple Stupid) principle and the Brewster Kahle archiving principle of keep everything you can but don't fret the details. Working at the Internet Archive we would boost that we were going for Quantity first, not Quality; the SAILDART data of 1998 is a pleasant past time since its Fixed Quantity becomes a lot easier to manage with each passing year.

9.2 DART formats Illustrated.

Six frames of 7-track tape supply a 36-bit PDP10 word

frame 1	frame 2	frame 3	frame 4	frame 5	frame 6
A A A A A A A	B B B B B B B	C C C C C C C	D D D D D D D	E E E E E E E	F F F F F F F
Bits 0 to 5	bits 6 to 11	bits 12 to 17	bits 18 to 23	bits 24 to 29	bits 30 to 35

Five frames of 9-track tape supply a 36-bit PDP10 word

frame 1	frame 2	frame 3	frame 4	frame 5
A A A A A A B B	B B B B C C C C	C C D D D D D D	E E E E E E F F	F F F F 0 0 0 0
Bits 0 to 7	bits 8 to 15	bits 16 to 23	bits 24 to 31	32 to 35 zeroes

7-bit SAIL ASCII to Unicode and UTF-8 table

	0	1	2	3	4	5	6	7
000	null	\downarrow	α	β	\wedge	\neg	ϵ	π
010	λ	$\backslash t$	$\backslash n$	$\backslash v$	$\backslash f$	$\backslash r$	∞	∂
020	\subset	\supset	\cap	\cup	\forall	\exists	\otimes	\leftrightarrow
030	$=$	\rightarrow	\sim	\neq	\leq	\geq	\equiv	\vee
040	\sim	!	"	#	\$	%	&	,
050	()	*	+	,	-	.	/
060	0	1	2	3	4	5	6	7
070	8	9	:	;	<	=	>	?
100	@	A	B	C	D	E	F	G
110	H	I	J	K	L	M	N	O
120	P	Q	R	S	T	U	V	W
130	X	Y	Z	[\backslash]	\uparrow	\leftarrow
140	'	a	b	c	d	e	f	g
150	h	i	j	k	l	m	n	o
160	p	q	r	s	t	u	v	w
170	x	y	z	{		ALT	}	BS

6-bit ASCII minus 040 code table

DART tape HEAD and TAIL record format.

word	name	value	description
0.	Type_Size	000006_000013	
1.	DART		sixbit/DART ω /
2.	BOT_EOT		sixbit/*HEAD*/ sixbit/*TAIL*/
3.	date_time		
4.	ppn		sixbit/DMPSYS/
5.	Class2Tape	XWD 2,Tape#	
6.	Rel_Abs		
7.	feet		
8.	words8	0	0
9.	minus1	-1	-1
10.	word10	0	0
11.	checksum		Rotated

DART file START and CONTINUE record format.

The MCOPY version#6 DART file (start and continue) record format has five parts.

	words	name	description
I	2.	TypeSize	DART record Type and Size
II	16.	RIB	WAITS File System Retrieval-Info-Block
III	18.	Leader	MCOPY extra baggage
IV	$0 \leq N \leq 10240. - 61.$	Payload	portion of the actual file data
V	23.	PRMERR	Previous Media Errors

Diagram of parts I and II

octal	decimal	symbolic	value	comment
...	0	type_size	(-3 or 0),size	size is 2 short of record length
...	1	dsk_or_error	'DSK'	constant
000	2	DDNAM	'filnam'	file name
001	3	DDEXT	XWD 'ext',Date	create (c)Date
002	4	DDPRO	prot, mode, time, date	write (m)Date Time
003	5	DDPPN	XWD 'prj','prg'	project programmer
004	6	DDLOC	track#	disk track
005	7	DDLNG	file length	PDP10 words
006	8	DREFTM	reference date time	(a)Date Time
007	9	DDMPTM	(T or P)dump date	(d)Date
010	10	DGRP1R	=1	first group
011	11	DNXTGP	=0	next group
012	12	DSATID	03164236 then 'RSK' or 'TSK' or 0	Storage Allocation Table ID
013	13	DQINFO	=0	defective 154 times
014	14	zero14	=0	defective 32 times
015	15	wrtool	'program'	write program name
016	16	DDWPPN	XWD 'prj','prg'	write project programmer
017	17	DDOFFS	=1	

Diagram of parts III, IV and V

octal	decimal	symbolic	comment
022	18	_DART_	sixbit/DART_/_
023	19	File_Con	sixbit/*FILE*/ or sixbit/CON_#/_
024	20	date-time	when MCOPY reel written
025	21	MC_SYS	sixbit/_MCSYS/_
026	22	two_reel	XWD class=2 and MCOPY reel#
027	23	one_one	XWD 1 and 1
030	24	Feet	MCOPY reel position
031	25	0	
032	26	-1	
033	27	0	
034	28	Words_To_Go	payload words remaining in file
035	29	0	
036	30	0	
037	31	0	
040	32	0	
041	33	0	
042	34	0	
043	35	0	
044	36...	file blob data payload	
000	-23	PRMERR	0
001	-22	"	0
002	-21	"	0
	...	"	0
024	-3	"	0
025	-2	'\$PEND\$'	046045_564404
026	-1	checksum	XOR

SAIL file system metadata

DART database files

9.3 Recipe for cooking DART data into GNU/Linux files.

Mount a large disk on /data or logically link to your large disk or /data/2014

Read Time Capsule into /dartrecords

De-compress and De-tar the time capsule files

Setup destination pathnames for undart-2014.

All destination pathnames are optional, if undart can not access a path it will not output files of that format. If there are NO destination pathnames found, undart will stderr a brief usage message and exit 0 for success. Setup destination pathnames for the GNU/Linux relinking command scripts. Setup destination pathnames for the HTML generation and relinking command scripts.

Compile and run UNDART-2014

Convert DART tape records into octal text and database CSV metadata of the SAIL files. Explain MD5 hashing of data blobs and serial numbering of the data blobs. The DART metadata is preserved in database CSV lingua franca.

Create a SAIL database and load the CSV metadata files into tables.

Convert 1-to-1 into current presentation format.

SAIL files into current presentation open source formats. Unix like UTF8 file systems, tar files and for web publication HTML5, CSS3, SVG, PNG, OOG.

Join fragments into comprehensible objects.

Merge pieces that go together.

Provide handling for multi version (sequential temporal) documents to show each point in time, or a best final and/or comprehensive copy.

Provide handling for multi version (cut-and-paste derivative work) documents.

Remove the ancient redundancy which has served its purpose of transmitting a message to us.

Redact damaged data (single byte drop out to zero, and such)

Static scholarship.

Talmudic scholarship and historical commentary about the meaning and significance of SAIL content.

Dynamic re-enactment.

Theatrical performance of scripts, Re-enactment, historical simulation, running models of the Look-n-Feel of life at the 1st SAIL.

End of preview. Now for a long saga concern the 1st Exegesis.



Figure 9.1: Innermost Matryoshka Doll

9.4 Program text of /home/sail/undart-2014.c

9.5 Program text of DART.FAI[TAP,REG]

9.6 Program text of DSKUSE.FAI[ACT,REG]

9.7 Remove DART packaging from SAIL file data

The first step in SAILDART exegesis is the removal of the DART outer wrappers from the content it carries. The content of SAILDART is Signal, the DART format is a Carrier. There is some noise, 63 DART record gaps, XOR checksum failures. There are no *rotating checksum* failures of the 60 byte HEAD-TAIL records. and the DART-v6 'ERROR.ERR[ERR,OR]' logging of yet earlier tape media defects, which all here simply bundled into tagged blobs along with all the more seemingly valid SAIL-WAITS file objects.

In the Carl Sagan book and movie, titled 'Contact', a message is received from Extra Terrestrials, with sufficient in-band clues (computer programs and documentation) to help the reader fully decode the whole alien message; which turns out to be instructions to build a machine (which I believe becomes in the story an instance of the extra terrestrial itself — but then the movie ending leaves one free to fill in the blanks as one wishes). So too to a lesser extent the message of SAILDART decodes itself and requires building a machine to fully instantiate itself.

One of Professor John McCarthy's A.I. challenge problems, was to write a program for a robot to travel from Palo Alto to Timbuktu. John perhaps expected something to do with Micro PLANNER or the Advice Taker enu-

merating practical steps like a travel agent booking a sequence of airplane reservations. John did not like our proposal (Baumgart, Moravec and possibly other A.I. lab Star Trekkies) that the robot transmit its blueprints, software, sufficient wired funds and technical instructions over a modem to Timbuktu to recreate an exact replica of itself at the African site. It is odd to reflect that I still recall John giving a rambling lecture on applying mathematical proof of correctness to travel bookings for an airline with one flight and one seat.

Starting with a naive vague idea about PDP-10 tape encoding and formats (and some personal experience, mostly forgotten, of the technology in that time period, and some email from Marty Frost) it has been possible to resurrect the in-band DART data records, the DART source text, the original DART documentation, the DART executable, and to convert the large DART database files.

So here we can see a dying empire, a major university computer research system, and its academic community suffering yet further disk failures (and heroic recoveries) and writing journals and messages like a monastic community during plague years of what has gone missing and what has survived on each disk crash and each budget reduction.

The final messages, system alert notices, and documents include a wake in "hope of the resurrection".

The second major task is converting ancient digital content into modern formats.

For some individuals this has often been experienced as a personal crisis as commercial vendors

ruinously lock-in their customers to proprietary formats that either disappear without a trace or force the users to buy upgrades of the format without long term recovery mechanism for old content.

Like digging out and unwrapping (or x-ray through) a Pharaoh mummy, inside a sarcophagus, inside a tomb.

Third task, redact damaged content.

The fourth task, clean the Augean stable, i.e. reduce redundancy. Redact exact clone copies.

Diff / Merge / GIT the incremental changes, with special attention for the append-only journal and log files.

Break into time stamped message snippets, dump into CSV (Comma String Value) files for database tables,

index by date-time, from who, to whom, about what. Then make a best effort attempt

to search, count, join, merge, sort and display results on the web.

Distinguish content which was publicly funded and published at the time from private communications and personal files.

DART magnetic tape defects The DART records have spans of garbled bytes. All files that DART deemed worthy to write to the permanent tapes where written twice. So most files appear two (or more times) within the dart records with the exact same bitwise content. The first example of a single missed dart tape defective byte that I found was rather obvious, a little to the

left of the center (at the base of the neck), in a much studied very early digital image. The good copy is at

[http://www.saildart.org/N.DAT\[XAP,BGB\]1](http://www.saildart.org/N.DAT[XAP,BGB]1)

Nearly as good, but with the defective black zero byte at row col is

[http://www.saildart.org/N.DAT\[XAP,BGB\]2](http://www.saildart.org/N.DAT[XAP,BGB]2)

So for a well curated SAILDART collection, only the first copy is included.
FAIL symbol table format.

9.8 SAIL text into UNICODE UTF-8.

9.9 Disassembling DMP executable binary into text.

9.10 SAIL software highlighting into cross-referenced HTML.

Seven bit text encoding at SAIL. As tabular illustration.

Arrows 001 down arrow↓ becomes Unicode u2193 represented in UTF8 as \342\206\223

136 up arrow↑ u2191 and \342\206\221

137 left arrow ← u2190 and \342\206\220

027 double arrow ↔ u2194 and \342\206\224

031 right arrow → u2192 and \342\206\222

Greek letters SAIL codes 002 003 006 010 007 are α β ε λ π become Unicode u03b1 u03b2 u03b5 u03bb u03c0.

In ASCII 007 is the BELL. When I see pi characters π π π on a display terminal I hear TTY bell dings.

Logic and Math SAIL codes 004 037 024 025 005 026 and 016 017
 $\wedge \vee \forall \exists \neg \otimes \infty \partial$ become Unicode u2227 u2228 u2200 u2203 u00AC u2297
and u221E u2202

Relations and Horseshoes SAIL codes 033 034 035 036 and 020 021
022 023
 $\neq \leq \geq \equiv \subset \supset \cap \cup$ become Unicode u2260 u2264 u2265 u2261 and u2282
u2283 u2229 u222A

9.11 Exercises

1. Exercise One: Assuming that you are a deputy SAILDART digital curator, expand the SAIL-DART-2014 time capsule SD chip into a large disk as database, web site and file system.

2. Exercise Two: Report back to me on the details concerning your exercise #9.1 work; or just wait for the 2015 edition.
3. Exercise Three: Assuming you are a lover of deep trivia, I have left the analysis of the GAP and the ERROR blobs for you, unless I get to it in 2015 or 2016.

Chapter 10

Meta Data - Taxonomy

10

Taxonomy is the classification of the SAIL objects by DART metadata, by SAIL file system metadata, by their apparent content type (Magic type or Mime type) and by what are now called On-Page Attributes.

10.1 SAIL file system metadata

Filename, date time written, file protection, file mode and the file length in PDP-10 words.

10.2 DART tape record metadata

DART metadata from the MCOPY run

Low density tape# and record#

High density tape# and record#

Latter day DART segmentation of the MCOPY

Byte offset in the 1998 DART byte vector : 0xAABBCC

dart segments (low density records + 63 gaps + 43 short skips)

10.3 Text verses Binary

The major text editor was named “E”. Text files generated using “E” always have an ASCII page table at the front of the file embedded as a comment, and so look like this:

```
COMMENT VALID 00002 PAGES C REC PAGE DESCRIPTION C00001
00001 C00002 00002 C00003 ENDMK C;
```

10.4 Role of the Digital Curator

10.5 SAILDART cooked metadata

Pub = True or False.

Here ‘Pub’ means that the SAILDART web server will serve the item. This is a manual classification based on latter day SAILDART policy implemented to determine which files belonged to which private individuals based on date spans and PPN codes. Files that are inappropriate to distribute on the internet in 2014 are marked Pub=False. However many files are marked Pub=True because they were very visible during the period 1972 to 1990 both at terminals at Stanford facilities as well as via telephone modem and the nascent ARPA network which became the TCP/IP internet; or because the known author has been contacted and has released the files as Pub=True for unrestricted web serving. Finally, I wish to note that Pub=False is a rather weak security classification, perhaps like the US government security level FOUO (For Official Use Only) in that the material is available as a collection to academic archival study. There is a handful of time capsule copies distributed around the world which are not encrypted.

Redacted = True or False.

Very few items are on a manual redact list. Hundreds of thousands of dart entries are redacted because they are exact duplicates in name, date, owner, protection and content. SAIL files were intentionally written twice as DART policy for the permanent MCOPY tapes.

Type. Text verses Binary.

The SAIL computer system encoded text in a non-standard 7-bit character code similar to ASCII. Furthermore the SAIL custom keyboards had extra shift modes (META and TOP) as well as two special interrupt keys named CALL and BREAK. And ALT was a character not a shift mode. End of line was marked by <Carriage Return><Line Feed>. There are even places with null padding between the <CR> and the <LF> so that the mechanical teletype head has enough time to slew from the right to the left side of a page.

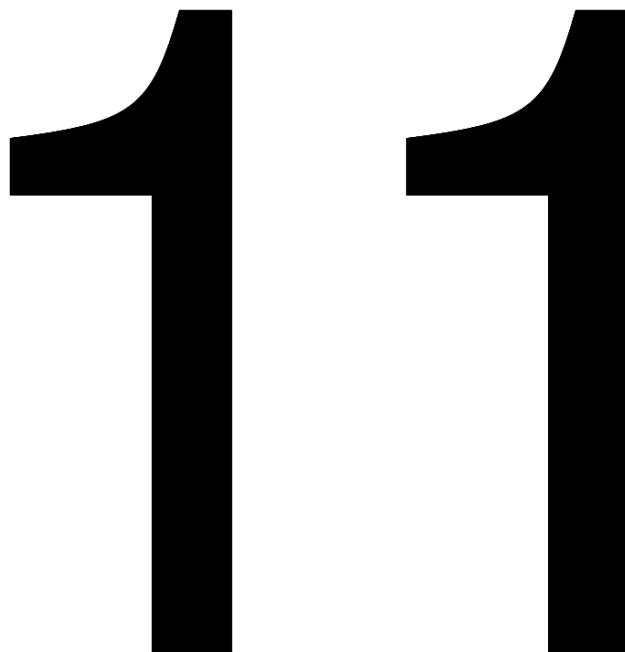
While first converting the heap of SAIL files it initially seemed obvious that they could be split into Text files and Binary files. This is ultimately misleading since computer software has both source text files and binary executable files. Likewise a piece of music has a source text and a binary representation. And the academic papers and technical documentation are written in mark up language source files (e.g. PUB and T_EX) that have to be processed to generate printed copy.

Copyright Status.**10.6 Exercises**

1. one
2. two
3. three

Chapter 11

Corpora Data Sets - dot • codes



Here a *corpus* is a set of files associated with a dot filename extension code. The plural of *corpus* is *corpora*. If you are like Les Earnest and think the word *corpora* is too Harvard snotty, you may say *corpuses*, which sounds to me like a mass murder. Sets of files, *corpora*, found within SAILDART are described. I shall postpone general remarks until after exhibiting specific • EXT labeled corpora.

11.1 • DMP

Dump files are binary executable PDP-10 machine code. Metaphorically, the computer was like a gravel truck in which the programs were first LOADED, driven around for a while, and then DUMPED. At DEC, but not Stanford, the metaphor morphed to dot SAV for *saved*.¹

As you may recall, the first word of a dump file loads into memory address 000074 of user space and the start address of the program is taken from the right half of 000140. Many of the dot DMP files have a symbol table for the debugging tools, DDT and RAID. I have used these symbol tables to disassemble the PDP-10 code and to joining the DMP files back upstream (against the current, often with a headwind and poor visibility) to their source files. In all there are 32730 DMP files of which 7428 have symbol tables. The largest and most interesting DMP files are built from many small source files, control files and separately maintained library packages. Software documentation does exist, reading ground truth in the source code and the disassembly listings is less arduous for the past since the quantity is orders of magnitude less than at present.

11.2 • FAI and • {nothing} and • MAC

The assembly machine code files number in the tens of thousands with three extensions: .FAI, .MAC and .{nothing}. Because the dot nothing files can be anything, but they are often FAIL, I have written a set of “Izzy” filters, for example isAssembly, isFail, isMacro, isLISP, isSAIL, isPASCAL and so on. (Then for a really obscure hacker story, Google for “Gosperism Soup” there is this Bill Gosper hacker dictionary story, re split pea soup. So Fail-P? SAIL-P? LISP-P? but I was a renegade so writing is-Foo rather than Foo-p.)

11.3 • LSP and • LAP

The LISP source code files have the extension dot LSP. The greater LISP family of associated languages and systems include PLANNER, Micro Planner, Met-alisp, mathematics (REDUCE and MATLAB), theorem provers and program verification systems.

11.4 • PUB and • TEX and • DOC

The pregnancy and birthing of digital typography, digital printing and desktop publishing occurred at Stanford in parallel with Xerox PARC, CMU and MIT (slightly earlier at Information International Inc, a bit later at Adobe, Imagen, HP and overseas) is documented inside the SAILDART.

¹There are DMP to SAV and SAV to DMP converters available as a bridge between SAIL and DEC for the emulation enthusiast.

11.5 • MSG

Back in the 20th century, *blogs* were called *bulletin boards* and the messages of a discussion group were appended to dot MSG files. Ordinary email also resides in dot MSG files. This 2014 SAIL archive shall attempt to keep personal and personnel messages in the dark for another 86 years. All the files of the SAILDART may be published at the stroke of midnight PST going into New Years Day Friday 1 January 2100. Never the less, Earnest and I (Baumgart) wish to make available the SAIL bulletin boards that were published on the “ARPA/Internet” in the 1970s and 1980s. Within each message file, each message in SAIL is prefixed with the partial differential symbol, ∂ prefix are often found in non .MSG files.

11.6 • SAI

Source code, written in the ALGOL like language named SAIL, comprises the .SAI corpus.

11.7 • DRW

The Stanford digital electronic design CAD software was named SUDS for Stanford University Drawing System. SUDS – Stanford University Drawing System – the suite of electronic design drawing programs with cryptic names D, PC, RPC, L, TD, LR and TRD.

11.8 • XGP

11.9 • PLT

11.10 • DAT

The dot DAT extension was used for generic binary data. Several long lived system programs wrote to dot DAT files with one or another convention for aggregating data by day, month and year; or at irregular intervals into dot OLD or dot ARC.

Every 15 seconds for over 18 years, the program named ACCT, alias JOB-NAM *SPY*, appended compute usage meter readings to its dot DAT file for of each day. ACCT made additional appropriate log entries for crash / reboot cycles and for when a user did a login or a logout. *Mirabula Dictu*, we can now view who the ACCT program saw as logged in at SAIL for almost every hour in the eighteen year period.

11.11 • FNT fonts

There are 4034 files with the extension FNT in the SAILDART. One gaudy example is my Bocklin knock off named BUCK75. On linux, you may fetch an octal dump of this font with a command like

```
wget -q http://www.saildart.org/BUCK75.FNT[XGP,SYS]_octal
```

The format for *.FNT[XGP,SYS] is

The early FNT

COMMENT STANFORD FONT FILE FORMAT.

WORDS 0-177:

XWD CHARACTER_WIDTH,CHARACTER_ADDRESS

WORDS 200-237:

CHARACTER_SET_NUMBER

HEIGHT

MAX_WIDTH (IN BITS)

BASE LINE (BITS FROM TOP OF CHARACTER)

WORDS 240-377:

ASCIZ/FONT DESCRIPTION/

REMAINDER OF FILE:

EACH CHARACTER:

CHARACTER_CODE,,WORD_COUNT+2

ROWS_FROM_TOP,,DATA_ROW_COUNT

BLOCK WORD_COUNT

For details concerning the early XGP

see Ted Panofsky's HM[H,DOC] section 18

aka SAIL Operating Note 56

titled Facility Manual,

by Ted Panofsky.

The latter day version of Ted's manual is at FACIL.TED[H,DOC]

[http://www.saildart.org/XGPSER\[J17,SYS\]](http://www.saildart.org/XGPSER[J17,SYS])

re BDF

http://en.wikipedia.org/wiki/Glyph_Bitmap_Distribution_Format

bdf2opcf BUCK75.BDF > buck75.pcf

cd /home/font

mkfontdir .

Add to X11 font path

xset fp+ /home/font

View font path

xset -q

tell X server to rescan the fonts

xset fp rehash

view font name

xfd -fn buck75

11.12 Generic • TXT • LOG • LST • OLD • TMP

The already mentioned the {nothing} file extension is the most numerous generic extension.

11.13 • DMD

Of the top dozen extension codes DMD was the first I did not recognize on sight. It is in the top rank because of LCS, Leland Smith. I leave it as an exercise to find out what it stood for and what the few people LYN, GHB, MMM, UW, DGP, BRP, PW, RD, SEK, MFB and MRC were using this extension for. Likely something to do with Music.

11.14 • REL

The REL extension is for relocatable files which were the intermediate assembler or compiler step prior to the loader, but again LCS and MUS have first and second place in terms of their file count with REL extension, while SYS and 3 are third and fourth in having REL files. My BGB sixty REL files in DART all seem to be part of GEOMED and might exist on DART since they were parts of shared libraries at times.

11.15 Languages • F4 • FOR • PAS • ADA • C • H

I was surprised to see how many Fortran, Pascal and 'C' programming language files exist in the SAILDART. There are 7360 files with .F4 extension, 1129 files with .FOR extension including the famous Adventure Game. There are 9074 Pascal .PAS files, ADA has 1972 files and 1356 .C files with 488 .H files. PL1 exists at 159 files.

The two files with CPP extension are text having to do with a Child Phonology Project. It could have been possible for C++ files to exist on SAIL. C++ was developed in the 1980s and had exploded in popularity before the final DART reel.

11.16 GEOMED files • B3D • CAM • CRE

Well this is *my* memoir, I am pleased that such a large set of geometric modeling related files has been preserved by DART.

11.17 SUDS Computer Aided Design Corpus**11.18 Digital Images Corpus**

The digital image formats •PIC •PIX •PIK •VID •DAT

11.19 Audio Sound Corpus

11.20 Extension EXT code Theory

The dot extension postfix to file names was a user option and not a mandatory MIME code for either the file system or the operating system. Over time various software packages developed that enforced EXT naming conventions.

11.21 Large, Medium and Small collections

In handling SAILDART, a further set of sets are named by T-Shirt size Large, Medium and Small. My typical mount point for a full set of SAIL objects is /Large which is for Curator access only. The /Medium is a comprehensive collection but with privacy filtering, copyright restrictions, redundancy removal, damaged data redaction and some relevancy redaction. For example, the DART tables are not included in /Medium, because which tape held which file name is part of the envelope not the message. The obsessive future scholar can go read the /Lcorpus. The /Small size S-corpus again has samples of everything but after extensive editorial selection - in particular the S-corpus attempts to have the latest or the best or even a typical version of each document. Ephemeral files as well as seven hundred ephemeral people are redacted. The ephemeral people are students, guest users, no name user codes, as well as users who left nothing but a trivial 'hello world' practice exercise or a few boilerplate files copied from elsewhere.

An important aspect of the Medium and Small collections is that they have been manually curated, best efforts, to protect personal privacy and to avoid copyright issues, and so can be widely mirrored and distributed. *If in doubt, leave it out.*

11.22 Document Sampler

Let me recommend that you read, or that you at least know about, the following particular SAILDART documents:

RESO.LES[UP,DOC] This is the SU-AI entry for the ARPAnet Resource Handbook as of 30 September 1977. It is a concise description of SAIL, the software and the documentation. Miraculously restored in HTML all the links pointing to old SAIL filenames are clickable.

WORKS.MSG[UP,DOC] A blog from 1981 to 1983 discussing work stations, which was then a niche market bigger and better than home personal computers. Indeed for us privileged few, the SUN work stations were our personal home computers.

YUMYUM[P,DOC] The San Francisco Bay area electronic restaurant guide with patron reviews. YUMYUM was the YELP for the decade 1973 to 1984. My hardcopy version of this is marked copyright reserved.

SYSTEM.MTG[A,REG] Minutes from the month SYSTEM meeting from May 1974 to October 1979. From this 185 page document it is easy to glean dates when major hardware, software or personnel changes were made. The KL10 arrived 1976-03-31. Amusing to quote, both the Librascope and Ralph Gorin were decommissioned on 1976-11-1:

PEOPLE

Wizard: Jeff Rubin is taking over as Chief System Wizard as Ralph Gorin goes looking for LOTS more trouble.

EQUIPMENT

Librascope: Rest in peace. Decided that the maintenance effort required is no longer worth the performance gain. We will either give it away or scrap it.

11.23 Exercises

1. Pick an EXT data type from this chapter and write the definitive guide to its content. Or merely add a few more paragraphs with illustrated examples to the above sections.
2. Write the T-Shirt sub sections for exactly what is in Large, Medium and Small at the end of 2015 and 2016. Coordinate the descriptions with the Exegesis chapter-9.
3. Append 48 more sampler document descriptions, then consider going for 100 total or so. Arrange into categories.

Chapter 12

Non-DART Collections



The SAILDART web site includes non-DART material relevant to the first SAIL. A Non-DART collection is simply archival items that were not on the DART tapes. Most such records are directly related to SAIL, however a few collections of records were entrusted to me which I could not turn away (for example the Bell Lab Unix newsletters for 1972 to 1990). The non-DART includes video of the reunion talks, digitized films and photographs, scanned

documents hardcopy of the PDP-6 and PDP-10 manuals Monitor and UUO manuals FAIL and PUB manuals and the Yum yum – online restaurant guide.

12.1 Scanned papers and books

12.2 Scanned photographs

12.3 Digitized SAIL film

12.4 Post DART video

12.5 Inventory of SAIL documents

12.6 Inventory of SAIL photographs

12.7 Inventory of physical SAIL objects

At the Baumgart residence in Los Gatos: A door key to the D. C. Power building, one reel of actual DART tape for the S1 user disk pack, and three ball point pens made of red and white plastic with a Stanford University Seal and teeth marks from a caffeine addicted graduate student. Paper based documents are listed else where (books, manuals, listings, notebooks, and photographs).

SAIL objects on display at the Wm Gates Building on the Stanford Campus in Palo Alto include the gold arm, the blue arm, a librascope disk platter, a keyboard, and a few others bits and pieces.

SAIL objects at the Computer History Museum in Mountain View include the ORM and the CART and very little else that is not a film or a document.

12.8 Exercises

1. Avoid mission creep. Simply hand off copies of the SAILDART for inclusion in larger collections maintained by computer history museums, university libraries, corporations, and individuals; rather than attempting to include their material inside SAILDART.
2. Ask people for their photographs, film, notebooks to scan and to catalog, DEC tapes, DAEMON tapes, even paper tape from SAIL if anyone of you can find any of it. I am less avid in collecting A. I. and computer science department material that does not directly appear inside the SAIL file system, that is a chore for other projects, libraries, museums, archives and individuals.
3. Write more in this chapter pointing out how the best of the non-DART material relates to the SAIL files. For example, where is the software that was used in each film? Where are the source files for each scanned paper ? Some folk have old paper listings they have saved from their

deceased academic mentor for so long that they now are able to ignore the fact that I have tens of digital copies of what that famous computer scientist wrote before, including that exact sacred piece of paper listing, and several versions from before and after ! OK, I see human passion now and then, so I lent out my big expensive flat bed scanner and we scan that paper listing and the Adobe OCR matches < almost exact? eh... close enough > the SAILDART digital copies. So the retired computer science professor clutching the paper listing returns to his beautiful home in the hills and as well returns to the illusion that only he has the key to what was written, and surely that is so (which of us here in the flatlands can really understand a couple hundred of pages of LISP in a week or two) ?

4. Archaeological approach to SAIL. I prefer keeping bits not atoms from the past, however I would be interested in helping anyone who wishes to investigate exactly what happened to the D. C. Power Building and its contents. PLEASE start with the Portal Pastures sign: which may be a new layer that is bolted on top of the old Stanford A. I. sign.

Part III

Containers for SAILDART

Part III Containers for SAILDART**PART**

chapter 13 Preservation

14 Publication

15 Participation

chapter 16 WWW static

17 WWW dynamic

18 WWW future



*Cicero used alliteration to push power points.
Preservation, Publication and Participation
are in chapters 13, 14 and 15;
the World Wide Web is in chapters 16, 17 and 18.*



Chapter 13

Preservation

13

The solution to archival preservation is simply to make copies of the message, store them in safe places that are accessible to readers by providing mechanism to assist reading the message and be sure to tell your peers, heirs and successors about your archive. Do not fret or linger at the archive. Go forth and do something completely different.

13.1 Disk storage at SAIL

The first disk hardware at SAIL was the infamous Librascope, followed by the IBM 2314, then then IBM 3330, which in turn was replaced with Ampex DEC RP06 disk drives.

13.2 Preserving SAIL files

Writing the original SAIL file system content from disk to DART tapes went on for 18 years. Copying from 3000 low density tapes to 229 high density tapes took 4 months, May to August 1990. Copying 50 Gigabytes from 229 reels of tapes to 25 Gigabytes¹ of disk in 1998 took less than four weeks of part time effort. Copying the 10 Gigabyte compressed file disk to a new SATA disk took two minutes twenty seconds just now, 2014-03-06, on an ordinary desktop Linux work station (and a further 96 seconds real-time to verify the md5sum). Making copies of this archive is not a major expense anymore.

13.3 Preservation Strategies

Time Capsules and Markers.

Carrying the Baton – Passing the Baton.

Carved in Stone. The quest for indelible permanent media.

Mention Stewart Brand and the Clock of the Long Now. Archive preservation is a one-way (half duplex) communication with intellectual entities who might exist in the far future (here on Earth or way out there in space).

1. Narrow casting. Time Capsules. Space capsule plaques. Write your message on permanent media and bury it in the desert or bury it in a library.

2. Broadcasting. Write your message on cheap media, make zillions of copies to send out in every direction.

3. "Baton passing" like in a relay race. Teach/Preach your message to your children and anyone else who might listen or read.

4. Chain letters that read: "This is the word of god. Thou shalt make many exact copies. It is the law. Do not change a single letter or dot. You shall be greatly rewarded after you die. If you do not copy this message or if you make any changes, then you will be severely punished. Amen."

I tend to favor baton passing methods and the more subtle kind of chain letter. Broadcasting and making time capsules can be fun too, but that is not my primary activity. Direct contribution to archive preservation work is done in the here and now. At an archive the past is yours to study and to re-live, while the future is yours for plans and predictions.

Indirect contribution seems a bit sad. Old people stand around at exhibits or re-unions to tell how their personal story has not been told correctly or is missing from the exhibit. Worse yet we tend to video tape the more articulate

¹50 to 25 by tar compress tgz

old people at these gatherings which increases the number of low grade gigabytes of digital history while the mere megabytes of higher grade digital ore may go unexamined and even unpreserved.

You do read (and act on) instructions from people in the past. Some people intentionally write instructions, orders, warnings, legal documents (last wills, living trusts and so on) and forward dated letters of credit to future entities (un-named direct descendant's as individuals, or particular groups of people such as a scholarship funds for aboriginal Tasmanians).

There is a book, about how to label a nuclear waste dump, Deep Time by Gregory Benford, cite Benford [1999], which relates that a dire warning surrounded with skulls, radiation warnings and bio hazards icons; will likely be read by future entities as "*dig here for the treasure*" which in the case of a time capsule might be what you most want to achieve. So some of your time capsule markers might just as well list extreme penalties. Curse of the Pharaoh.

13.4 Passing the Baton

13.5 Exercises

1. LOCKSS – Lots Of Copies Keeps Stuff Safe.
2. Mirrors.
3. Digital Library Software.
4. Estimate the cost of materials and labor for maintaining the DART backup tapes during the first 18 years of its service which started late in 1972. Then the cost of the final MCOPY tape-to-tape effort of 1990. Then estimating the virtual cost of the 1998 copy on the back of an envelope begins as 60+ hours of volunteer tape wrangling, and assumptions like: let's "pay" these wranglers \$100. per hour for their volunteered time, estimate the cost of the media they deployed at IBM Almaden to receive the copy pursuant of IBM goals to research related to data mining and massive near line storage systems.

year	'72:1980:'89	1990	1998	2016
media	3000+ tapes	229 reels	8x 9 GB disks	64 GB SD chip
cost of media			\$250. each	
labor hours	4000+	1000+	60+	0.1
cost of labor	\$25	\$50	\$100 per hour	\$100
total in the day			\$8000.	
new:old \$ ratio	2.88	1.81	1.45	1.00
present \$ Total				

Chapter 14

Publication

14

The container for SAILDART must not be opaque. Librarians expect visitors and circulation. Archives permit access. Depositories and Time Capsules are built to be read. Publication includes dynamic presentation and performance as well as static hardcopy media.

Exercises

1. Resume Friday afternoon office hours for SAILDART (starting at 3 or 4pm) in or near the present day A. I. Lab's Fishbowl conference room #120 Wm. Gates building on the Stanford Campus.
2. Start holding code review readings for the SAIL 1974 re-enactment software.
3. Self publish this Prolegomenon with an end-of-year ISBN edition. Reprint sets of the DEC and SAIL systems programming manuals need for 1974 hacking.

Chapter 15

Participation

15

Participation in the SAILDART Archive work is nearly zero, but not insignificant. The individual authors of the original files number fewer than 2000. The major institutions involved in creating SAIL files include Stanford University, DARPA née ARPA, NSF, NASA and so number fewer than ten. Except for those named below all the former SAIL people and their institutions are not involved with the SAILDART Archive. The work of setting up the SAILDART

Archive and running it for sixteen years, 1998 – 2014, has taken me I would first guess at least one *mythical* man year, 2000 hours, of part time effort. But that guess may over count time I have spent as a *user* of the SAILDART rather than as its *maintainer*.

15.1 Form of Organization: Karass.

In the science fiction book titled Cats Cradle, the author, Kurt Vonnegut, defined the terms *karass* and *granfalloons*. I see now that SAILDART is like a *karass*, a few individuals who occasionally coordinate to do something significant. Other more formally organized entities such as Stanford University, the Computer History Museum, the Internet Archive and IBM Research are more like *granfalloons*.

15.2 Participation Role: Digital Curator.

The 2014 SAILDART Archive Curators.

Next Contact Plan

It is inexpensive to make a compress copy of the original 52 Gigabytes so that anyone who I knew from my days at SAIL and who I trust enough to follow John McCarthy's verbal security directions gets a complete copy of the original with a tar ball snapshots of my undart Exegesis software for recent years. Most would be SAILDART archive assistants want BGB to do chores X, Y and Z and to whack the data into their Museum, Shrine, Library, Book, Web site or educational entertainment event. And for two SAILDART related events, I have ended up doing the video editing, youtube uploads and getting transcripts into HTML5.

For me (*as a computer science degree holder and garage based hacker*) an archive, like a library, has users as well as archivists.

The exact distinction between a time capsule, record repository, archive, library, museum and a web server is resolved by looking at information transfer transactions. The usage policy (both de facto and de jure) the nature of the agents to information transfer transactions.

For this simplification, a Time Capsule is a Write-Once-Read-Once message in an attempt to communicate over a period of time that exceeds the transmitters lifetime. The crash recorder on an airplane would be a time capsule.

The classic analog answer is that Museums focus on objects, Libraries handle books and Archives deal with records. Naive legal definitions are found on loc.gov In the digital world objects, books and records are all construed as information.

On the digital frontier,

The strong do what they can, the weak suffer what they must –
Thucydides.

15.3 SAILDART Institute Job Descriptions

At the Fantasy SAILDART Institute, there are many career hats to wear, and candidates for the various positions to be considered. As enumerated here a staff of 30 to 50 would suffice, with stipends of say \$120K per annum, this would be a \$5 Megabuck per year operation.

Title: Leader (de facto B.G.Baumgart recruited by John McCarthy, encouraged a couple of times by Don Knuth, tolerated by Frost, Gorin, Earnest and Reddy.)

Duties: Preserve the data from the DART tapes. And meet (or find) users, members, customers, clients and patrons for the project. 1. Define tasks, goals and subgoals in writing. 2. Set priority for the goals, suggest a schedule of what to work on next. 3. Assign resources to work towards each goal (time, equipment, personnel). 4. Review progress towards goals, then check success or failure per time period and budget.

Title: co-Leader (de facto Les Earnest; perhaps Frost, Selker, Gorin could be recruited as a standby co-leader; unfortunately the default responsibility might fall on my legal heirs and successors Leona, Stephen and David Baumgart.)

Duties: 1. Talk to leader about every aspect of the project; and 2. Substitute for (or become, or find) a new leader when the need arises.

Title: GNU/Linux systems administrator (de facto BGB, here and for every further role all the way down this section).

Title: Computer operations engineer

Title: Work station and laptop wrangler

Duties: purchase, integrate, install and maintain multi screen work stations. Provide staff with a plethora of compute power, I/O devices for storage, printing and scanning.

Title: Computer security expert (candidates: Whit Diffie, Ron Rivest and Bruce Schneier)

Title: Computer security operations monitor

Title: JavaScript programmer

Title: Python programmer

Title: 'C' programmers with PDP-10 experience (candidates: REG, PMP, TVR, BH, DEK, Sampson, Lampson, Supnick) also consider 'D' and 'C++' skilled candidates.

Title: LISP programmers (candidates: Allen, McJones, Hearn, Masinter, Dave Smith, Costello)

Title: Curator
 Title: Censor
 Title: Technical writer (candidate: DEK)
 Title: Technical writers for each specialty area (hardware, LISP, OS, Robotics, CG, CV, MTC, etc)
 Title: Proofreaders
 Title: Editor
 Title: Publication tool operator
 Title: Database administrators
 Title: Web site creative design
 Title: Web site content (HTML5, CSS3, PNG, SVG, OOG) open source tool expert
 Title: Web server admin (for Apache2 at the moment, consider adding Lighttp and Nginx)
 Title: Web analysis (for both internal logs and external links)
 Title: Legal expert for handling copyright issues (Creative Commons, Lawrence Lessig).
 Title: Comedy writers (candidates: Susan McCarthy and Tom Costello).
 Title: Catering for all the workshops and meetings.
 Title: Gourmet Chef
 Title: Sommelier and Oenologist (a Diffie / Earnest time share perhaps).

15.4 People I wish to contact

When the cookbook recipes for the chapters titled Exegesis and Static WWW are ready to be tested, contact Paul McJones, Dan Hartwig then Marty Frost all of whom have expressed interest in expanding a SAILDART time capsule into a file system and static web site to access records they will chose to integrate with the existing historical material at the Computer History Museum in Mountain View, the Stanford CSD, and the Stanford University Green Library.

When SYSTEM.DMP[J17,SYS] is running past LOGIN on the 'C' and 'D' emulators, get the implementation visible on Github and contact Ralph Gorin and Brian Harvey.

Tovar would be useful on Music, and possibly Foonly, Fonts and PDP-10 assembly code. I trust Tovar might be able to verify my implementation of the Dave Poole KA10 XCT AC \neq 0 opcode, since he wrote the firmware for it on the latter day Foonly models.

Bill Gosper was perhaps the first SAILDART user to find a lost theorem.

Bill Pitts concerning Space War, Art Samuels and checkers; as well as the early cyberculture of hacking into the lab on a night when the door was wide open.

Ted Panofsky concerning hardware.

Tom Gafford concerning SUDS and Foonly as well as the 1975 home computer "Maytag" PDP-11 disk channels.

Steve Russell concerning Spacewar and the pre-DART PDP-6 Lab, and even the Pre-Earnest A.I. Days.

When I have time to resume working on details of specific privacy issues, contact the Family McCarthy (ELF,SMC,) with respect to curating, releasing for public display, the JMC files on SAILDART.

Whit Diffie and Mary Fischer.

When the PUB emulator is working, first contact Larry Tesler and then all the others who have requested this or that PUB document as a PDF.

When ancient versions of T_EX are working, Art Keller and Les Lampert may be interested.

When SUDS is working, the Seattle Museum Group would be interested. They claim to have a 'C' version of SUDS but are lacking some supporting material and verification.

The 2014 circle of LISP people include Tony Hearn, Larry Masinter, and Paul McJones.

Les Earnest has often requested to see a running PARRY reincarnated, one recent PARRY correspondent who is attempting to find in the SAILDART a set of PARRY source code as well as the project code HMF (Higher Mental Functions) related papers and notes.

Large quantities of digital music related files do exist on SAILDART with respect to John Chowning(MUS), Leland Smith (LCS), Andy Moorer(JAM), Jim Gray and the many students who did the Stanford 206 Course work using the SAIL computer system. However, the largest audio files were kept "offline" on user disk packs or even written to analog magnetic tape for performance, and so are not in the DART specific data of the SAILDART Archive.

The SAIL programming language, Dan Swinehart and Bob Sproul

S1

Sun and Cisco

DEC related people: Supnik, Lampson,

15.5 Exercises - this time for the Reader.

1. Exercise One: Write to me a wish list of items to add and ways to improve the static SAILDART web site.
2. Exercise Two: Describe in writing (and then even build) the version of a SAILDART institution you would like to see.
3. Exercise Three: Use the SAILDART public material in some other project.

Chapter 16

Static WWW

16

On a static web site ever URL request retrieves a static flat file from a disk file system with no generation of page detailing or template substitutions (other than SSI, Server Side Includes) or database retrievals. In the case of the static SAILDART web site this mechanism has been taken to the extreme.

2012 note

The <http://www.saildart.org> is a static web site with an Apache2 server. It lives inside a file system container called the Large Corpus (see T-shirt sizes) mounted at /Lcorpus on a multi disk raid mirror.

2013 note

It is now April 2013, and I need to write down a brief description of the 2012 SAILDART bird-in-hand so I can replay the conversion batch scripts and make a new improved static web site as well as generating more database table CSV files for dabbling with dynamic site implementations and to find interesting content for general publication, for one to one communication, for personal insight, or for redaction to the far future when all the first generation biologically carbon based human SAIL people are dead.

2014 note

Yet another cycle of the static web site will be attempted to provide improved pagination, source code highlighting and a lot more internal anchor cross reference links.

2015 note

Changed the top page to include slots for the re-enactment, reunions and the annual draft copy of this book as a PDF.

2016 note

Drastic reduction of the set of pages on the web site by removing the individual password protected user areas. It has been over a decade, this functionality is nearly worthless.

2017 note

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce ac sagittis velit, egestas fringilla tellus. Vestibulum aliquam lacus accumsan, sodales mauris a, fringilla mauris.

Exercises

1. Do something useful with the Apache logs.
2. Replace Apache with Lighttp and/or Nginx.
3. Add style sheets and markup for source code coloring.

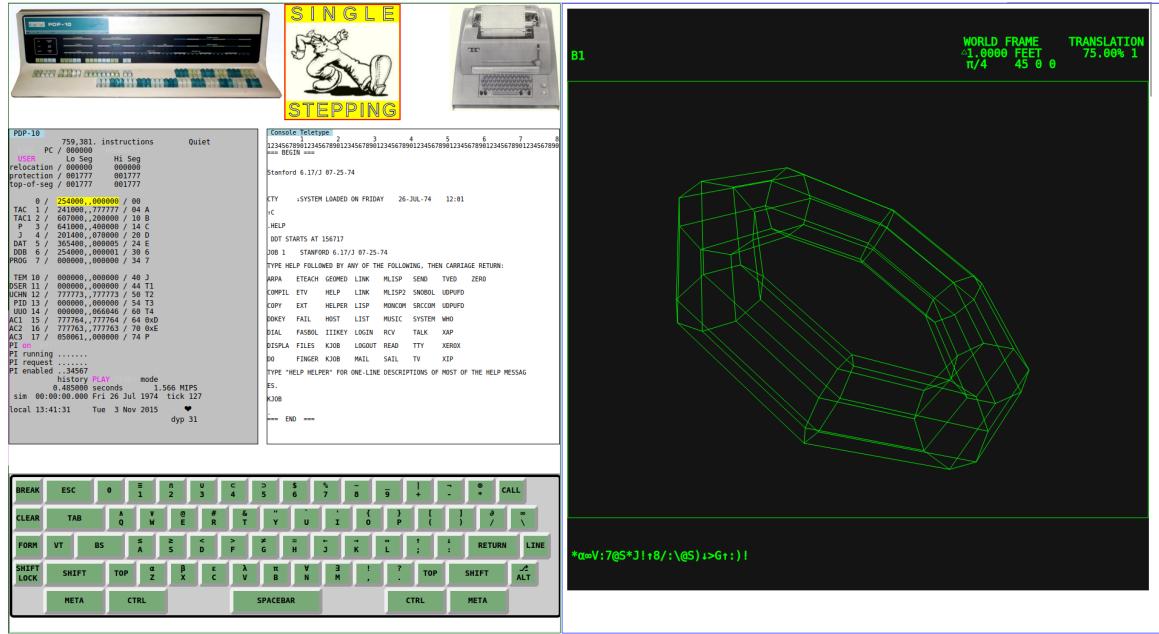
Chapter 17

SAIL System ReEnactment

17

This chapter reports the work in progress on my quest to run all the old SAIL software via the SAILDART.org web site. Like a 19th century railroad steam engine, or an 18th century tall masted sailing ship, running the 1974 SAIL mid 20th century time sharing system is a major stunt that I have not adequately achieved yet.

**Screen Shot of KA10 running SYSTEM 1974
Javascript inside a Chrome browser**



PDP-10 KA : Good PDP-10 simulators are available and PDP-10 hardware does exists in museums in running order. However, running the SAIL operating system of 1974 has proven to be difficult because that particular PDP-10 machine, a model KA-10, had unique modifications and the best emulators target later models such as the KL-10 which differs substantially in I/O and memory mapping / relocation. For example, at Stanford the non-user mode execute XCT instruction with a non-zero AC field was modified (perhaps by Dave Poole) to do peek or poke into user space.

I/O devices : Each and every major I/O device was a one of a kind. With extreme pride, our beloved pioneers hacked out each device. The supporting documentation as well as the actual artifacts are gone from my horizon. The year 1974 is a portal. If I can get 1974 running again in simulation or as a re-enactment - then the following decade and a half becomes much more visible because the documents and digital drawings at the foundation of the 1970s start at SAIL with SUDS, PUB, TeX and lead directly into Xerox, Apple, Adobe, Autodesk, Cisco, Sun and other places. In the case of Sun and Cisco, the relationship is direct as well as sinister; but the law suits have been settled, and the personal relationships can be viewed from a distance. Let the muse of History, Clio, be the judge. She need not report for another hundred years or more, lets say 2114 or 2214.

Non ASCII : The character set was a NON standard version of ASCII. The XGP Xerox Graphics Printer was one of the first three in the world. The LPT line printer was a custom drum. The disk channels were built by Stanford personnel. The real time digital calendar clock was designed and built by Phil Petit and was only one of many interesting pieces of digital hacking mounted in the Kludge Bay. The III vector display processor was the only one of its kind every built, likewise the first large disk the Librascope disk, the interface to the IBM-2314 and IBM-3330 were hacker built home brew, the Television camera A to D was built by Dave Poole who also decided one day to swap the character code for ALT and Right Curly Brace because that was the way it was on a raw model-35 Teletype. Standard ASCII the '}' is octal 0175 to this day, at SAIL for twenty five years the octal 0175 was the ALT character 'r' and the right curly was at octal 0176.

Alan Kay in 1968 showed people a cardboard mockup of his Dyna Book idea; those of us who saw it were skeptical and said such things like "OK and someday a PDP-10 will fit in a suitcase". Ken Olson (Digital Corp CEO) did not like having a big machine product line.

The PDP-10 Javascript simulation of GEOMED, the 1974 SYSTEM console teletype and 1974 SYSTEM single user III vector display are close to being ready for 2014 previewing on the public web site. PUB, POX, T_EX and XGP web output are a high priority.

PDP-10 history

PDP-10 fandom

Hardware Difficulties

Simulation of ideoyncratic hardware : Given that the design drawings for the 1974 generation of the hardware is lost, the only path forward for me is a careful review (a trace) of executing the operating system and patch out all the devices which are not interesting to me and making up fictional representations of the devices I care about based on the interface visible in the code itself augmented with a few paper documents I saved, polaroid photographs I took and memories I can recall. That includes the special SAIL keyboards, the III vector display, the XGP printer, the LPT printer, the CTY console teletype, and the IBM disk interface.

Perhaps SAIL hardware built after 1974 will reveal itself via SUDS drawings if and when I get SUDS running again. Post 1974 SAIL, SUN, Cisco, DEC, S-1, Foonly and so on designed their early digital wonders at SAIL using the SUDS design software.

Verification of ancient system software : The system programmers who wrote and debugged the SAIL system used program listings and built-in debug software name XDDT for executive mode DDT. Mechanical toggle switches

and small incandescent lights (light emitting diodes, LEDs, in the late sixties, were expensive and only glowed a dim red) on the console also implemented Examine, Deposit, Stop, Start. These listings were generated by the assembler and printed on wide paper. On the left side of the listing page there were octal columns showing the address and value generated for each line of PDP-10 source code shown on the right side of the page. There were also printed symbol cross reference tables called CREF. Since the listings, cref and obj could all be rebuilt from the source these derivative files were intentionally NOT preserved on the DART backup tapes. One might bootstrap into this past world by getting a version of the FAIL assembler to work, or by disassembling the binary DMP files and trying to join matching lines of disassembly to the corresponding source lines. Such latter day neo assembly and matching has to deal with the extensive use of conditional assembly and macro expansion.

Model KA-10 Priority Interrupt : The early models of the PDP-10 simply implemented the PDP-6 seven level interrupt mechanism which was actually inferior to the late models of PDP-1 which had sixteen level sequence break hardware for scientific real-time data collection.

PDP-10 Code Reading : The major part of getting the old SAIL system running has involved single stepping the code, then setting break pointers to isolate the unknown portions.

17.1 SAILON-76

After a gap of 38 years, here is
 Stanford Artificial Intelligence Laboratory Operating Note
 number 76 seventy six.

17.2 Keyboard

```

left1: break escape
row 1: 0 1 2 3 4 5 6 7 8 9 + - *
top 1: □ ≡ ∩ ∪ ⊂ ⊃ $ % ⊗ _ | ↵ ⊗
right: call

left2: clear tab
row 2: Q W E R T Y U I O P ( ) / \
top 2: ^ V @ # & " ' { } [ ] ∂ ∞

left3: form vt bs
row 3: A S D F G H J K L ; :
top 3: ≤ ≥ < > ≠ = ← → ↔ ↑ ↓
right: return line
  
```

brk	esc	0	\equiv	\cap	\cup	\subset	\supset	\$	%	\sim	$-$		\neg	\otimes	
clear	tab	Q	\wedge	\vee	@	#	&	"	'	{	}	[]	∂	∞
form	vt	bs	\leq	\geq	<	>	\neq	=	\leftarrow	\rightarrow	\leftrightarrow	\uparrow	\downarrow		
shift	shift	top	α	β	ϵ	λ	π	\forall	\exists	!	?			return	line
lock	Z	X	C	V	B	N	M	,	.	top		shift		alt	
	meta	ctrl			space					ctrl		meta			

Table 17.1: keyboard layout

```

left4: shift_lock shift top
row 4: Z X C V B N M , .
top 4:  $\alpha$   $\beta$   $\epsilon$   $\lambda$   $\pi$   $\forall$   $\exists$  ! ?
right: top shift alt

row 5: meta ctrl space ctrl meta

```

17.3 Exercises

1. one
2. two
3. three

Chapter 18

Road-map Atlas

18

Prolegomenon to SAILDART finished at the end of Chapter 17. Chapter 18 is *Postlegomenon*, it is my SAILDART project To-Do List. For comic effect in the next paragraph, I italicize project management buzzwords, which like rocks in my stream of consciousness, always pop out as I raft down my *action items*.

I divide the project into *tasks* and *milestones*, requiring *tools* and *skills*. Some tools will be brand new and others ancient, as needed to work in the SAILDART archive. Then I implement an appropriate Strategy, Tactics and Logistics for this hobby project. Then review again the phases illustrated by the Hindu trinity of Brahma, Vishnu and Shiva to Create, to Maintain and to Delete; always “*Knowing what to throw away, knowing what to keep*”. Then to study the horizon for for *Threats and Opportunities*; and to write down a Wish List, a To Do List, and a Business Plan, and to quote from Brooks’ MMM and from de Marco’s Peopleware, to spread out the sheets of costs, fixed and variable, to evaluate my ROI, Return on Investment, fiscal and psychological, to write a Calendar, time line, and to do the maintenance chores and repairs. I am looking forward to my Exit-Strategy, Completion, Hand-Off and Closure. I intentional omit *monetization*. Finally, *going-forward* (by-day) is the *bottom-line*.

18.1 My Goal Immediate

Finish the chapter-17 emulator(s) then finish writing this introductory book. Print a limited circulation edition each year without a satisfactory chapter-17 and with too many short place holding paragraphs (like this one) that are only slightly better than *Lorem Ipsum*.

18.2 Tools and Skills

Learn how to use all the Adobe Creative Suite as well as all of TeX, Latex, LyX and TeX packages for the LyX Memoir Class.

18.3 Patrols and Sorties

A patrol map is for a defensive exercise. A planned route is revisited at intervals or upon an alarm to do maintenance, updates, and to look for threats and opportunities that can be seen by peering out between the parapets of your castle. A sortie is an offensive maneuver, to sally forth into unknown territory, the sortie maps are brief suggesting the new direction, the immediate few obstacles and the distantly perceived target area.

18.4 Maps

- Prolegomenon Table of Contents is a Grid Iron Map of the SAILDART playing Field.
- Map for going from DART into 64-bit words, UTF-8, SQL, HTML, SVG, PNG and OGG.
- Map for validating a PDP-10KA with I/O devices to run the SAIL time sharing system J17 SYS

- Map for publication
- Map for renewing tools, skills and domestic/office environment.

18.5 Priority: Decision verses Recreation

For me, the secret to finishing my PhD thesis in graduate school was to stop programming and to just do the writing. Like finishing an art project in Kindergarten one must stop kneading the wet clay and form some recognizable object to take home to your parents.

A Quantum Physicist was hiking alone in the woods and came to a fork in the trail *<brief pause>* then he took both paths. That is what I do when it comes to programming languages, since so many new ones look so attractive, I clearly must try them all. That is recreational programming. It is not the way to complete a work to show to others.

18.6 Exercises

1. Execute on sections 18.1 and 18.2 and 18.3
2. Implement an exit strategy and get on to doing other projects and writing other books.
3. Take a hike. *Since I run a monthly hiking group, this is already part of my routine.*

Back of Book sections

Table of Video

I have uploaded to Youtube videos on the McCarthy celebration page,

Wikipedia Articles

The Wikipedia explicitly states that it is an encyclopedia and as a work of academic writing the Wikipedia is a Tertiary Source of knowledge. The ground rules for contributing to the Wikipedia exclude first person writing and original research material. The content of the SAILDART archive is a primary source, the text here in this Prolegomenon is secondary source material.^f

Fonly Computer design project at SAIL that spawned SUDS and spun out the company named Fonly which designed and sold a series of PDP-10 clones.

WAITs Latter day name for the operating system at SAIL, formerly called SYSTEM which had been derived from the Digital Corporation PDP-6 operating system called Monitor.

Spacewar! The computer game that is a bright shinny thread running across the fabric of late 20th century computer culture. Where the bright shinny metaphorical glow is a P28 green phosphor.

Table of Scanned Documents

Digital Picture Gallery

Analog Picture Gallery

Bibliography

Gregory Benford. *Deep time : how humanity communicates across millennia.* Avon, New York, 1999. ISBN 9780380975372.

John Markoff. *What the Dormouse Said: How the 60s Counterculture Shaped the Personal Computer.* Viking Adult, 2005. ISBN 0670033820. URL <http://www.amazon.com/What-Dormouse-Said-Counterculture-Personal/dp/0670033820%3FSubscriptionId%3D0JYN1NVW651KCA56C102%26tag%3Dtechkie-20%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0670033820>.

Trademark Attributions

I am still looking at the Digital Equipment Corporation debris field for a U.S. commercial entity claiming trademark for PDP® and DEC® perhaps it is somewhere in a drawer at HP (did HP just fission again?) or at one of the 1990s spin offs. Ask Gordon Bell at MS SF research concerning DEC and PDP ?

Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries.

UNIX® is a registered trademark of The Open Group.

PDP® might belong to Hewlett Packard.

DEC® might belong to Hewlett Packard.

Photography Credits

Baumgart. I took most of the early analog pictures (using both Polaroid and Kodak cameras) and I still have hardcopy prints for further scanning and to prove my ownership. I also took a lot of the early digital pictures. I gladly give credit to KRD, (K. Randy Davis) for our featured digital image of Annie Liebovitz ! The Liebovitz oeuvre is a copyright troll mine field .. I shall just risk including the pictures of myself that she took in 1972, her legal successors might forge my granting her permission.

Glossary**Spell Checking**

name spelling
Baumgart coinages
computer jargon
abbreviations

Full Word Index, Concordance**Name Roster, Nomenclatura****Index**

Index

- Arthur L. Samuels, 47
Brand, Stewart, 27
Crispen, Mark (MRC), 29
Earnest,Les(LES), 49
Frost, Marty(ME), 51
Gafford,Tom(TAG), 49
Gorin, Ralph (REG), 81
Helliwell, Richard(Dick), 22
John McCarthy, 47
Kay, Alan, 151
Knuth, Donald Erwin (DEK), 51
Les Earnest, 47
Liebowitz, Annie, 27
McCarthy, John, 20
Moorer, James Anderson (JAM), 28
Russell, Steve, 21
Tovar(TVR), 49

Colophon

For this edition I have been using the open source document authoring tool \L^XYX built above Don Knuth \T^EX and Les Lamport \L^AT^EX. My 2009 to 2016 platforms include GNU/Linux work stations, Apple OS-X BSD laptops and one *Cadillac* Mac Pro workstation.

Write a colophon as complete as that in Wolfram's A New Kind of Science.

The End

"Euclid's Elements has no preface or introduction, no statement of objectives; it offers no motivation or commentary. It opens abruptly with a list of 23 'Definitions' at the beginning of Book I."

- Richard J. Trudeau, pg23 'The Non-Euclidean Revolution'

This concludes the 2016 edition of the Prolegomenon to the SAILDART Archive. This document is a marker (a treasure map) pointing at a time capsule (the treasure). This document is an introduction to what is in that archive. Please go use the SAILDART Archive. You will be rewarded !

Amen.