Lawrence Berkeley National Laboratory
Recent Work

Title

A REVIEW OF THE MAJOR PROBLEMS ENCOUNTERED DURING TEN YEARS OF OPERATING A LARGE SCIENTIFIC COMPUTING SYSTEM

Permalink

https://escholarship.org/uc/item/8bd399wz

Author

Osborne, Carol F.

Publication Date

1970-06-01

eScholarship.org Powered by the California Digital Library
University of California

Conference on Automation of Scientific
Researches on the Computer Base,
Novosibirsk, USSR, September 14-18, 1970

UCRL-19842
Preprint

c.2

A REVIEW OF THE MAJOR PROBLEMS ENCOUNTERED DURING TEN YEARS OF OPERATING A LARGE SCIENTIFIC COMPUTING SYSTEM

RECEIVED
LAWRENCE
RADIATION LABORATORY

SEP 25 1970

LIBRARY AND
DOCUMENTS SECTION

Carol F. Osborne

June 22, 1970

AEC Contract No. W-7405-eng-48

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 5545

LAWRENCE RADIATION LABORATORY
UNIVERSITY of CALIFORNIA BERKELEY

UCRL-19842

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

A REVIEW OF THE MAJOR PROBLEMS ENCOUNTERED DURING TEN YEARS OF OPERATING A LARGE SCIENTIFIC COMPUTING SYSTEM

Carol F. Osborne

Lawrence Radiation Laboratory
University of California
Berkeley, California

ABSTRACT

The Lawrence Radiation Laboratory, Berkeley, California has an automatic film scanning device attached on-line to an IBM 7094 II computer. The system monitor is multiprogrammed to coordinate computer use between three levels of priority: 1) real-time automatic scanning of bubble chamber film, 2) processing of the scanned data, and 3) batch processing of other analysis programs. Bubble chamber events are processed, completely automatically, at the rate of 150 per hour.

Operational problems encountered in growing to 500000 events per year will be discussed. These can be categorized as follows: 1) people and machine problems in manual (or semi-automatic) scanning, 2) maintaining communications between operating staff and programmers, 3) restart and recovery procedures, 4) quality control, 5) hardware problems, 6) time and event accounting procedures, and 7) data summarization.

Submitted to:
Conference on Automation of Scientific Researches on the Computer Base. September 14-18, 1970. Novosibirsk USSR.

Conference sponsored by:
The Computing Center and the Institute of Automation and Electrometry of the Siberian Department of the Academy of Sciences of the USSR.

Special problems arise because six scanning machines record asynchronous, independent data onto a single tape. It is often necessary to edit specific data from the scan tape, thus there are routines to process data from a specified experiment, start and end processing on a specified date and time, and ignore specified data. Some of the edit routines will correct minor hardware problems which occur between periodic maintenance of the scan machines. Some error correction is built into the scan machines to allow the scanners to delete tracks, views, and frames in which they know about errors. Rescanning can then be done immediately.

Hardware problems.

Hardware problems are a whole class by themselves. One would like to think that one's computer would always function perfectly and that periodic running of diagnostics would detect and periodic maintenance would correct all problems. Unfortunately it is not this precise. On most subtle problems, maintenance people will always question a program and programmers will always question the hardware. There is always a bit of "if you would fix your computer" vs. "if you would fix your program", however by establishing good communications between the programmers, especially the system monitor programmers, and the hardware maintenance people, most hardware problems can be isolated and fixed relatively quickly. Both sides should have some understanding of the other's problems. Complaining that "the machine isn't working" and getting a reply that "all the diagnostic programs run all right" is really not getting the problem solved. The programmer or computer operator must analyze a problem in sufficient depth to provide adequate guidance for maintenance personnel. Good understanding on both sides of the hardware/software problem is useful so that special diagnostics can be written to help isolate a problem in the hardware. In one serious instance, we spent almost three weeks constructing special diagnostics to help the hardware people determine that one phase of an instruction in the disk commands was not working properly.

This same problem of hardware vs. software applies to the FSD hardware and programs. A diagnostic program that does fairly exhaustive testing of the FSD hardware is run before every long production run, an on-line summary stating the status of the hardware is printed, and an off-line listing of the data causing the error is written. This sounds quite logical and workable, however when both the FSD and the program which controls it were first tried, there were all sorts of problems in both hardware and software. It is much easier when you have either a proven program with which to test a new machine or visa versa than when both are new.

Operator problems.

Maintaining good communications between the operators and the programmers is even more important than between hardware and software personnel. Since we have our own computer and our own

INTRODUCTION

The bubble chamber data processing system at Berkeley is a mature (more than ten years old) and successful (over 1.8 million events processed) system. The solutions to the problems encountered in its development have made it successful.

BACKGROUND

Typical bubble chamber film processed is shown in Figure 1. Three views of each event are taken and typical experiments take from 30,000 to 300,000 events.

The process of achieving an automatic system has been a series of evolutionary steps, each of which has served to automate the most time consuming process in the chain of events. The progress can be measured by noting the key advances.

1. A series of programs to analyze data obtained by manual methods were written.
2. Semi-automatic data gathering from off-line machines was started with precision microscopes.
3. A flying spot digitizer (FSD) was developed to automatically scan areas of bubble chamber film which have been pre-scanned and selected by human scanners.
4. Fully automatic scanning of bubble chamber film without human intervention was started with a series of pattern recognition programs. Although no human scanners are required in the fully automatic mode, the system accomodates pre scanning of experiments with very sparse events.

PROBLEMS

Problems in the man/machine interface.

The largest single source of problems can be traced to human errors in pre scanning the film. Whether these errors are caused by carelessness, fatigue, or any other reason is somewhat immaterial; the conclusion is the same - humans are major sources of error. Currently 75 to 80 percent of the human errors show up as improperly located data points in the film scanning process. Some errors are caused by poor quality film which can not be digitized and the remainder of the errors are divided among many factors such as: the scanner entering the wrong experiment or roll number, improperly identifying the type of event, etc. We currently check for 27 types of human errors which have occurred in the past with sufficient regularity to warrant special corrective routines. The programs must be flexible enough to allow all sorts of special features - one rather interesting request was to allow 39 days in each month. (This was to maintain compatability with a different event identification system.)

special purpose system monitor, we can train operators to have a good understanding of the system. Normally we do not get terse notes from operators saying "the program failed" - they are much more thorough in their analysis of the failure.

Operators must be able to initiate restarts in a multiprogramming environment which contains three different levels of priority. The operator should be able to determine if the problem is in one of the levels, in the system monitor, or in the hardware and should attempt a restart. When all else fails, there are methods for obtaining summaries of the processing which has occurred. Obviously a great amount of effort has been placed on file protection to protect the system and the different levels.

The operators also keep log books of the processing, perform quality control, and handle production runs. When a system is initially started, many of the problems relating to a high production situation do not occur. However, as the work load increases, it is necessary to have successful restarts, detailed summaries, and precise communications between programmers and operators.

Quality control problems.

Quality control is partially handled by the operating staff, however the majority of all quality control (especially in a completely automatic system) is integrated into the system software. Quality control programs verify the number and type of events processed, the number of failures and reasons for failure, the rate of processing, and a general analysis of operational information. Programs also give the operator the current information necessary to initiate restarts.

Another portion of the system software is concerned with time accounting and event accounting. It is important to know where the time is spent in a large system in order to analyze efficiencies of programs, the monitor, and the operators. The system monitor produces time accounting information consisting of start, end, and used time and type of run, etc. for each program run. This is then analyzed and daily and monthly reports are produced which are used for operator supervision, time budgeting, and scheduling. Event accounting keeps track of all the events as they are processed through the whole series of analysis programs. It is concerned with the status and progress of the experiment. The status gives the user information about the events input and output to prevent any bias in the processing and the progress tells the operators what should be done to complete the experiment.

Data summarization problems.

One of the largest problems in any system is to summarize and present final data in a useful form. There is a logical compiler in the data summarization system which is used to select levels and categories of data. Data plots, histograms, and listings can then be obtained. Provision is made for manually selecting ambiguous events and it is also possible to do comparisons of hypotheses concerning these kinematically ambiguous events in the summary

programs. The largest problems in data summarization in a large system are knowing in what stage the data are at any particular time and in keeping the data properly identified.

SUMMARY

In summary, the problem in automating any system is one of efficiency, i.e., to be fast and accurate. The solution requires a knowledge of the system status and the ability to take corrective action on the problems as they occur. The primary solution is one of communication: people with people; scanner with scan machine; operator with computer and with the system; programmer with operator and with the computer; and administrator with the system and with the people.

Figure caption

1. Typical bubble chamber film is on the right. The automatic pattern recognition program results are in the middle and left.

References

1. H. S. White, "Status of the LRL Flying Spot Digitizer". Proceedings of the 1966 International Conference on Instrumentation for High Energy Physics, Stanford. (Clearing house for Federal Scientific and Technical Information MES, Springfield, Va.).

2. H. S. White and D. E. Hall, "Unassisted Analysis of Bubble Chamber Events". Meeting of the Society of Photographic Instrumentation Engineers, N. Y. June 9-10, 1969. UCRL 19210.

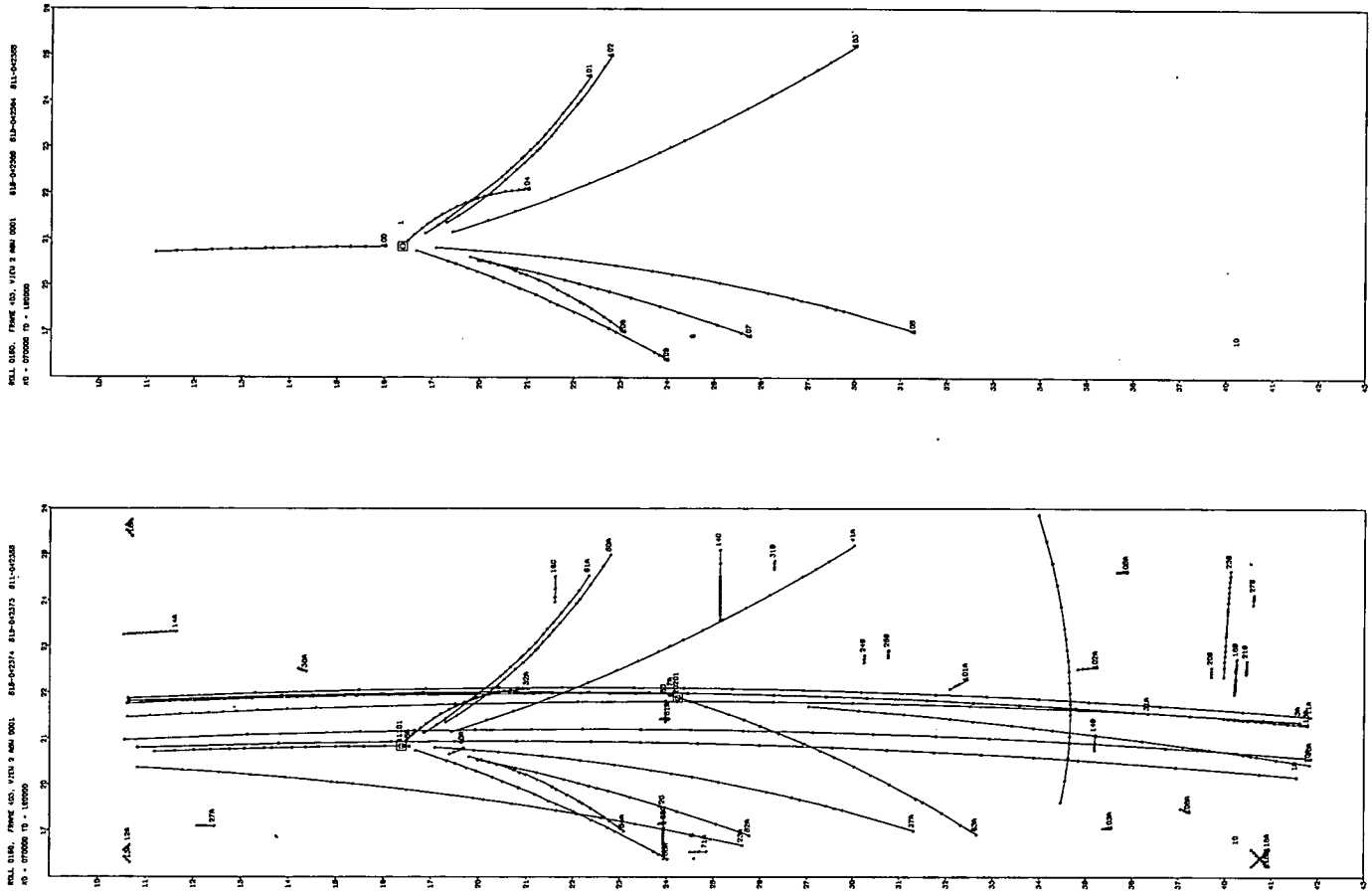
3. H. S. White et al, "DAPR: Digital Pattern Recognition Approaches Production". Proceedings of the International Conference on Advanced Data Processing for Bubble and Spark Chambers, Argonne, Ill. Oct 28-30, 1968. UCRL 18543.

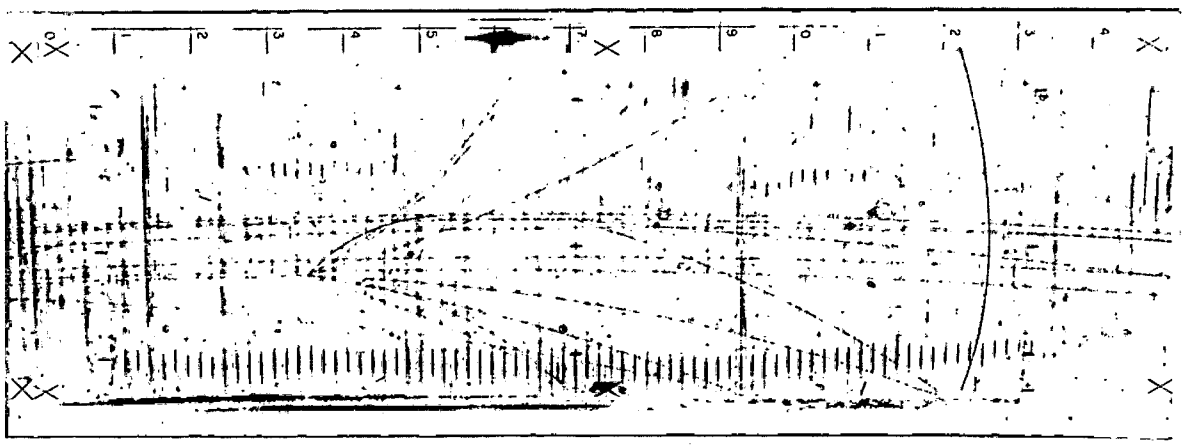
4. C. Osborne, "Operations On-Line to a Large Computer". Proceedings of the International Conference on Advanced Data Processing for Bubble and Spark Chambers, Argonne, Ill. Oct 28-30, 1968. UCRL 18547.

5. Loren Shalz, "FSD Hardware Monitoring". Proceedings of the International Conference on Advanced Data Processing for Bubble and Spark Chambers, Argonne, Ill. Oct 28-30, 1968. UCRL 18548.

Acknowledgements

This system was developed by the Data Handling Group directed by Howard S. White. The work has been sponsored by the U. S. Atomic Energy Commission.





XBB 6810-6155

Fig. 1

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

TECHNICAL INFORMATION DIVISION
LAWRENCE RADIATION LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720