Proceedings of the Eastern Joint Computer Conference

Held by the

Joint IRE-AIEE-ACM Computer Committee

Washington, D.C., December 8-10, 1953

Published by

The Institute of Radio Engineers, Inc.
1 East 79 Street, New York 21, N.Y.

Printed in The United States of America

Price $3.00

PROCEEDINGS OF THE EASTERN JOINT COMPUTER CONFERENCE

Theme: Information Processing Systems—Reliability and Requirements

PAPERS AND DISCUSSIONS PRESENTED AT THE
JOINT IRE-AIEE-ACM COMPUTER CONFERENCE,
WASHINGTON, D.C., DECEMBER 8-10, 1953

SPONSORS

The Institute of Radio Engineers Professional Group on Electronic Computers

The American Institute of Electrical Engineers Committee
on Computing Devices

The Association for Computing Machinery

Published by

THE INSTITUTE OF RADIO ENGINEERS, INC.

1 East 79 Street, New York 21, N.Y.

ADDITIONAL COPIES

Additional copies may be purchased from the following sponsoring societies at $3.00 per copy. Checks should be made payable to any one of the following societies:

L. G. CUMMING, Technical Secretary
Institute of Radio Engineers
1 E. 79 St., New York 21, N.Y.

R. S. GARDNER, Assistant Secretary
American Institute of Electrical Engineers
33 W. 39 St., New York 18, N.Y.

Association for Computing Machinery
2 East 63 Street, New York 21, N.Y.

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 54-9071

Conference Committees

JOINT IRE-AIEE-ACM COMPUTER COMMITTEE

Chairman: J. H. HOWARD, Burroughs Research Laboratory

Exhibits: PERRY CRAWFORD, JR., Intelectron Corporation

Finance: J. C. McPHERSON, International Business Machines

Publications: W. D. LEWIS, Bell Telephone Laboratories, Inc.

F. L. ALT, National Bureau of Standards

E. G. ANDREWS, Bell Telephone Laboratories, Inc.

M. M. ASTRAHAN, International Business Machines Engineering Laboratory

E. BROMBERG, Association for Computing Machinery

L. G. CUMMING, Institute of Radio Engineers, Inc.

S. H. DODD, Massachusetts Institute of Technology

R. S. GARDNER, American Institute of Electrical Engineers

G. G. HOBERG, Burroughs Research Activity

D. H. LEHMER, Institute for Numerical Analysis

F. J. MAGINNISS, General Electric Company

W. L. MARTIN, Telecomputing Corporation

G. D. McCANN, JR., Department of Electrical Engineering, California Institute of Technology

J. D. NOE, Stanford Research Institute

L. N. RIDENOUR, International Telemeter Corporation

J. H. WEINER, Eckert-Mauchly Computer Corporation

S. B. WILLIAMS, Consulting Electrical Engineer

TECHNICAL PROGRAM COMMITTEE

Chairman: HOWARD T. ENGSTROM, Remington Rand, Inc.

Assistant Chairman: ROY C. BRYANT, Remington Rand, Inc.

J. S. ANDERSON, Airinc

J. J. EACHUS, Department of Defense

M. W. SWANSON, Department of Defense

CONFERENCE COMMITTEES

LOCAL ARRANGEMENTS COMMITTEE

Chairman:	MARK SWANSON, Bureau of Ships, Department of the Navy
Assistant Chairman:	A. E. SMITH, Office of Naval Research, Department of the Navy
Finance:	RALPH I. COLE, Melpar, Inc.
Inspection Trips:	MARGARET R. FOX, National Bureau of Standards
Registration:	L. R. JOHNSON, Department of the Air Force
Publications:	H. N. LADEN, Bureau of Ships, Department of the Navy
Publicity:	G. W. PETRIE, International Business Machines
Reception and Hotel:	STANLEY F. REED, Reed Research, Inc.
Exhibits:	L. D. WHITELOCK, Bureau of Ships, Department of the Navy

S. N. ALEXANDER, National Bureau of Standards

ROY C. BRYANT, Remington Rand, Inc.

E. W. CANNON, George Washington University

C. A. HENSON, Bureau of Ships, Department of the Navy

J. J. A. JESSELL, Federal Power Commission

C. V. L. SMITH, Office of Naval Research, Department of the Navy

S. B. WILLIAMS, Consulting Engineer

PUBLICATIONS COMMITTEE

Chairman: W. D. LEWIS, Bell Telephone Laboratories

K. M. COLLINS, Bell Telephone Laboratories

J. R. HARRIS, Bell Telephone Laboratories

W. KEISTER, Bell Telephone Laboratories

J. H. McGUIGAN, Bell Telephone Laboratories

TABLE OF CONTENTS

Opening Address
Joint Computer Conference

JOHN HOWARD†

IN THE AUDIENCE this morning are many visitors who have come from all over the United States—and from many other countries as well—to attend this conference. Your response has been most gratifying; and on behalf of the Joint Computer Committee, I wish to welcome you to its Third Annual Eastern Conference.

Washington is a particularly favorable location for this conference for three major reasons: First, since it is one of the major computing centers in the country, it provides some unique opportunities for the inspection trips shown on page 9 of your program. Thanks to the co-operation of the cognizant agencies, this will be the first opportunity for many of us to visit these pioneer computer installations.

Second, Washington readily furnished the enthusiastic and capable leadership required to manage a conference of this magnitude.

In this connection, I would like to have you refer to page 10 of your program, which shows the membership of the Local Arrangement and Technical Program Committees. I will not read all their names but wish to give special mention to Mark Swanson. He has done an outstanding job in guiding the management of the conference. Also, special mention should be made of Dr. Lewis of the Bell Telephone Laboratory and his Publication Committee who are responsible for the publication of the Conference Proceedings.

Finally, Washington is the home of a number of pioneers in the computer field. Many of them have never received the public recognition they deserve because of the classified or administrative nature of their work.

To these men we owe a real debt of gratitude for their early and continued support in the development of the data processing field. Many of these men are responsible for making policy decisions as to future courses of action to be taken —now that the field has reached its present state of development. Some of these decisions are momentous and will have a direct bearing on the security of the nation.

For this reason, I am particularly pleased to extend special welcome to them to attend this conference and exhibition so that they may hear first hand our presentation on a very important theme—"Information Processing Systems—Reliability and Requirements."

† Research Activity, Burroughs Corporation, Philadelphia, Pa.

Keynote Address

HOWARD T. ENGSTROM†

Introduction

I SPEAK TO YOU this morning as Chairman of the Technical Program Committee for the AIEE-IRE-ACM Joint Computer Conference. At the first of these conferences, in December 1951, at Philadelphia, a program was devoted to a description of existing electronic computers or those under development. The second conference, at New York, in December 1952, was devoted to a review of input-and-output equipment necessary for efficient usage of computer systems.

In considering subject matter for the 1953 joint conference, your Committee chose the theme "Information Processing Systems—Reliability and Requirements." Perhaps some explanation of this theme is in order.

I have been exposed to the rapidly developing art of large-scale digital computation since 1941. It appears to me that the art has gone through a series of phases. Prior to 1941, developments in electronic counters had reached a point where counting of electrical pulses could be carried out at megacycle speeds. The advent of World War II, and the pressure for making use of these high-speed units in military equipment, resulted in a series of developments which made possible supplying information to electronic accumulators at speeds commensurate with their rate of operation.

These developments included memory systems, notably the magnetic drum, electrostatic-storage tubes, and acoustic-delay lines. Although the perfection of these memories did not take place until after the war, basic developments were carried out during the war years. Fundamental work also was carried along in logical structure of computation devices and the command system.

The third phase of development in the art was that discussed at the conference of December 1952; namely, the input-and-output equipment which made possible communications from the memories and computing elements to the external world. Within the last two years, therefore, it is felt that the electronic-computing art has achieved maturity in that we now have well-balanced equipments. That is to say, we have high-speed arithmetic elements; we have a hierarchy of memories which can feed these arithmetic elements; we have peripheral equipment, such as magnetic tapes, high-speed printers, conversion equipment—which maintain contact between these memories and the external world.

This maturity of the art also applies to the industrial structure behind it. Shortly after World War II there appeared to be a lack of interest in the art by the larger companies who were probably most capable of advancing it. Many small companies were started on the basis of government contracts for specialized computers. Many such companies did not survive the financial problems. In the last two years, the interest of larger companies in this new field has been clearly demonstrated. Many of the smaller companies have been absorbed into these larger structures so that industry is now in a position to produce these equipments as rapidly as the demand for them increases.

It appears to me that the phase of the art into which we are now entering is that of applications. In applying these computers to many problems, questions of reliability and dependability have been repeatedly raised. Your Program Committee has therefore arranged this conference in essentially three parts. The first group of papers consists of statements of requirements in a variety of applications. Requirements in weather prediction form a typical case of scientific calculations. The use of electronic computers in the life insurance business is illustrative of many applications in the accounting area. Computer applications in air-traffic control problems are in an area in which reliability and performance may be a matter of life or death. The real time-control systems are illustrative of applications in the direction of complete automation.

The second part of the program is a series of papers giving actual performance characteristics of existing computers. These papers present results of operation, in many cases over a several year period, and since the papers are presented in general by those responsible for producing computational results, they should give a good indication of the success of existing computers in meeting requirements in reliability and performance.

The third series of papers is descriptive of work carried out during the past several years in connection with improving reliability of components and will provide valuable information from the position of component manufacturers in insuring the dependability of electronic-processing systems.

In closing, let me repeat that I firmly believe electronic information-processing equipment to have arrived at a state of maturity, and the next phase in the art is one of applications. Many of the advances in the field of applications will be made by the users of this equipment. It is gratifying to have reports from a variety of prospective users at this conference.

There are, of course, many applications to military problems which, for security reasons, cannot be discussed at this open conference. We had hoped to plan a classified session, but were unable to complete the arrangements. I am sure you are familiar with military applications, however, through individual relationships. It is hoped this conference will bring together information concerning reliability and requirements sufficient to answer questions of many prospective users.

I should like to express my appreciation to Mr. John Howard and his Joint Computer Committee for the parent organizations, for their great assistance in drawing up this program.

† Engineering Research Associates Div., Remington Rand. Inc., Arlington, Va.

The RETMA Support of the 1950 Computer Conference—A Progress Report

THOMAS H. BRIGGS†

IN ATTENDING the numerous technical conventions and symposia that are held each year, we listen to many learned papers. Problems important to our daily work are thoroughly discussed. Generally we return home with ideas that will forward our own engineering investigations. But we often find the problems discussed are larger than the individual or organization can handle. With mounting technical complexity, these problems are more frequently encountered.

The problem of tube reliability is of this nature. It is complex; it affects more than one industry. It causes us to say, "I wish someone would do something to improve reliability." During the past few years many groups were formed that effectively attacked the problem of reliability. Today, tubes are damned far less often, and the performance of electronic equipment has improved.

In December, 1950, a Conference on Electron Tubes for Computers was held at Atlantic City. It was sponsored jointly by the AIEE, the IRE, and the Panel on Electron Tubes of the R&DB. Over 300 engineers attended. The technical success of the conference was due in large meaure to: Dr. C. V. L. Smith, of the Office of Naval Research, Computer Section; Dr. A. L. Samuel, of IBM and the Panel on Electron Tubes; and Dr. Mina Rees, also of the Office of Naval Research. The papers defined problems and informed the audience about actions taken in individual situations.

As we left Atlantic City it was apparent that the Conference had been a real success. From every side one heard: "What can be done? This problem is too big for our laboratory." There was spontaneous demand that means be provided for co-ordinating effort leading to improved tube reliability for computer applications. The object of this paper is to show that action was taken, and how it has progressed during the three years past. We hope that as similar needs arise from this and future conferences that there may be stimulation to employ similar logical solutions.

The Radio-Electronics-Television Manufacturers Association was selected as a logical area for such an activity. It has been a basic plan of RETMA to encourage the study of standardization of components within special levels of refinements whenever such a need is shown, even though standards exist for the same basic components used in other applications.

† Formerly with Burroughs Adding Machine Corp., Philadelphia, Pa.; now Electronics Consultant, Norristown, Pa.

In standardization, the initial studies are made by committees of engineers who are experts in that field. Therefore, for studying some special new class of components an entirely new group may be established. The story of activities in the field of electronic computers is typical of activities in promoting new products for a special class of equipment.

In May, 1951, after an organizing meeting was held under the auspices of the RETMA Engineering Department, the group decided that the electronic tube-reliability problems merited their undivided attention. Other components used in the computer field were left to other committees. Since the Joint Electron Tube Engineering Council (JETEC) is the agency that was set up by the two interested trade associations, RETMA and NEMA, to handle all problems relating to electron tubes, this Electronic Computer Committee was reassigned as a subcommittee of JETEC Committee JTC-5. This committee covers all receiver types of tubes. Thus the JTC-5.5 Sub-committee on Electronic Computer Tubes was created. Representation has consisted of tube producers, computer manufacturers, laboratories.

The problems apparent to JTC-5.5 at the start were:

1. To define reliability as required in the field of electronic computers.
2. To classify the special computer application needs, as they relate to tubes and crystals.
3. To determine the best test methods and the conditions to apply in making them.
4. To establish means for translating the conditions that are peculiar to computers into those that are useful to tube manufacturers for developing types especially suitable, and in controlling products.

The first concrete step was to prepare and distribute a lengthy questionnaire to all possible computer companies and laboratories. The gamut of questions were phrased to permit concise answers, generally of a "yes-no" nature. They concerned existing operating environmental conditions, types of tubes and crystals in preference of use, means for controlling and monitoring quality, and testing methods, all for use in this application.

An encouragingly high percentage of returns came back, well completed, for the major undertaking of correlation. These data were summarized and, in April 1952, were circulated to the participating groups for their information and use. All questionnaire returns were coded for maintenance of company anonymity with regard to answers.

The answers indicated that several new tube-characteristic ratings were needed to define properly the tubes and the rectifier-type crystals required by the computer industry. To accomplish this, three task forces were set up to develop new formats for the JETEC Rating Sheets that were to be used in registering new tube and crystal designations for these applications; the three classifications were:

1. Crystal Diodes,
2. Twin-Triode Tubes,
3. Multi-grid Tubes developed for computer service.

The crystal diode format was brought to near completion and recently has been transferred to JTC Committee-14.2 for final acceptance and future use.

Following six draft stages the twin-triode format has just been submitted to JTC-5 for adoption. The multi-grid format required the submission of a second questionnaire to industry, but it too is nearing completion. Returns received during the summer of 1953 have been collated for study.

The new formats for tube ratings embody the previous information, but contain supplementary controls for plate current, cut-off characteristics, pulse ratings, plate-current balance between sections, contact potential limits, and several dependent criteria.

Normally, time-dependent criteria have not been admissible in establishing tube ratings. For more conventional usages tubes are considered interchangeable if they operate alike for relatively brief periods of time. However, it is the consensus of both tube and computer engineers that the definition for interchangeability in computers must include operation over a lengthy time. Thus, interface impedance build-up and heater failure with "on-off" voltage cycling for a particular tube can only be evaluated by studying those conditions over extended periods of time.

Incorporation of the time-dependent parameters as rating criteria poses many problems and technicalities. The actual method for specifying those factors is now under study.

Actions of JTC-5.5, or any committee in fact, must be predicated upon potential economic values. These two tube formats, when finally developed, will then be submitted for adoption by the RETMA and NEMA as regular JETEC Standards. The Tube companies can then employ the special data forms to supply ratings on the tubes generally used by the computer industry or when rerating tubes applicable for this service. The use of these formats is expected to provide ample economic return for the efforts expended.

Another major project on the agenda is the preparation of an Application Manual, directed to tube producers and to computer-design engineers. The reasons which lead to the format and rating decisions will be stated and means for applying the tubes most effectively will be pointed out. Interpretations of ratings, and means for determining the effects of unusual operation conditions will be included. It is expected that this manual will be of great use to component engineers and to circuit designers.

JTC-5.5 is cooperating with other groups in their work. Liaison is being established with JTC-5.3 concerning the establishment of a suitable method, capable of standardization, for determining *short-circuits*. This will require a rigorous definition of a short-circuit, relating such factors as time, resistance, and degree of tapping force. The causes of short-circuits, when they are intermittent and of extremely short duration, are extremely difficult to determine, and hence must be eliminated by accurate and standardized tests.

Co-ordination with the Cathode Committee of the American Society of Testing Materials—designated B4-VIII-A—has been established. Their work on development and standardization of a method and conditions for *interface impedance* testing is aided by interlocking memberships from the two groups. This problem is probably one of the most important confronting the electronic computer industry today.

So far, the work has been confined to digital-computer vacuum tubes. There have been requests to extend this both to gas-discharge tubes and to storage tubes. Extension of the scope of JTC-5.5 to include these tubes has recently received authorization from JETEC Council and these new problems now await handling time on the part of the working members.

The computer industry is far flung geographically. It is essential that the ideas of the entire industry be represented in the discussions of the Computer Tube Sub-Committee. Consequently, it is planned to establish a West Coast Coordination group. In this manner the results of our deliberations should have greater scope and authenticity. The tube manufacturers should be able to place greater dependability upon the recommendations from this joint tube producer-consumer sub-committee.

Initially JTC-5.5 concerned itself with germanium diodes as well as tubes. Obvious close relationship in use, as well as similar reliability objectives indicated a natural "junction." In the spring of 1953, cognizance of the diode crystals was transferred to the Committee on Solid State Devices. The program started in JTC-5.5 is being continued by JTC-14, under Mr. J. R. Flegal of Bell Telephone Laboratories.

JTC-5.5 Committee Membership

Mention has been made that this progress has been due to the co-ordinated efforts of the engineers of tube manufacturing companies and design engineers of computer equipment.

Chairmanship of the Sub-committee JTC-5.5 is alternated yearly between tube producers and computer equipment manufacturers. The first chairman was

Mr. J. A. Goetz of IBM. His energetic enthusiasm and organizing abilities gave the group an excellent initial momentum. In the second year, Mr. R. E. Higgs of RCA was appointed chairman, and capably extended the progress made in the first year. The Sub-committee JTC-5.5 is responsible to Committee JTC-5 on Receiving Tubes, and thereby is accorded major support and assistance through the latter's Chairman, Mr. C. E. Coon of the Tung-Sol Electric Company.

Other Components for Electronic Computer Equipment

As the progress of JTC-5.5 became effective and the difficulties experienced by the computer industry with tubes became better understood, improvements were announced. Other components in the circuit then became the centers of inquiry by RETMA Committees set up for this purpose. Most of these inquiries relate to Reliability matters.

A number of problems were mentioned, and so an RETMA committee designated G-2 (Computer Components) was recently established to cover the special needs of products in this field.

There are many items that have found applications in electronic computers. Those used in digital computers are, for the most part, common types of components, but precision, freedom from drift under unusual ambients or over extended periods with aging, constitute special considerations. Also there are a number of other components of a very special type used, especially in the analogue computer field.

The G-2 committee acts as a special co-ordinating group that determines what special problems need attention, and writes up a brief, outlining each of them. The Executive Council of the RETMA Engineering Department uses these briefs to set up new projects in the particular committees involved, or even authorizes the establishment of a new committee if an appropriate committee is not on the present roster. In each case the objectives are clearly outlined, but G-2 committee will provide supplementary suggestions when required by the committee handling a particular project.

The first of these G-2 sub-groups is G-2.1, concerned with Precision Potentiometers, under the chairmanship of Mr. J. R. Altieri of Technology Instruments Co.

Long ago RETMA established committees on such common components as resistors and condensers. Thus, outlines of special computer requirements for these components will be prepared by G-2 and submitted to these existing committees for study and action.

Another group of interest to this computer conference is G-3 on Magnetic Amplifiers under the chairmanship of Mr. E. V. Wier of Magnetics Incorporated.

Receiving tube manufacturers are located mainly on the East Coast. The computer industry, although young, has manufacturers and development laboratories scattered from coast-to-coast. Further, there are analogue as well as digital computers.

It is the desire of JTC-5.5 to have their results as authentic and as applicable to the entire computer field as is possible. Tube producers have found that there is a broad market for tubes that are specially-rated according to JTC formats. A number of tubes have been registered by the RETMA Data Bureau for use in computer circuits. Effective co-ordination of computer requirements will result in greater applicability of these formats. Extensive use of the format and availability of tubes meeting the listed requirements will thereby be accelerated.

JTC-5.5 has therefore inaugurated a new procedure to overcome the problem of geographical separation, and to increase the authenticity of its results. Arrangements have been made to establish a branch committee on the Pacific Coast. Agenda for meetings will be exchanged so that programs, discussions, and work may proceed in parallel. Through interchange of ideas, desires, and results, it is expected that progress will be on a broad front, and the results should be more widely employed. The tube manufacturers will have more certain data for tube ratings. The computer industry and groups with similar tube interests, will have even greater expectancy of achieving reliability.

In closing, I have no doubt but that this conference will serve to high-light its own problems on Information Processing Systems. I sincerely hope that you, too, may be stimulated to establishing working means which will lead to the solution, or at least the amelioration, of the new crop of challenging problems.

Discussion

R. L. Carmichael (Bell Telephone Laboratories): How are these committee reports and findings made available to industry? How does one keep abreast of the developments of the various committees?

Mr. Briggs: This is a problem which is of universal interest these days because every hotel you go to, on any trip, you always find a convention taking place and generally of a technical nature. On several committees of which I have been a member, it has been our attempt to sell to the membership that they should carry the gospel back to their home organizations. We have also attempted through personal papers, such as this one, to keep the general field abreast of what is going on in the technical committees. Most technical committees are very glad to have guests attending any of their meetings. JTEC 5.5 sub-committee meetings and their task force meetings, which is where the work is generally conducted, are well documented in the monthly notices put out by the RETMA office in New York. I would suggest that any of you who are interested should determine the membership representation in your company, and get them to keep you abreast or write to the RETMA headquarters should you desire to become a member of any of their committees.

The Use of Electronic Data Processing Systems in the Life Insurance Business

M. E. DAVIS†

(*Presented for Mr. Davis by J. J. Finelli*)

Introduction

IT WAS UNDER the impetus of the last war that several rapid computing devices were developed. They did very difficult and important mathematical work for the government and, in doing so, demonstrated a capacity to handle work similar to that found in many offices. It is not surprising, then, that at the end of the war some of us began to feel that machinery of this kind might be very useful in the life insurance business.

A relatively large amount of recordkeeping, computation, and analysis, is required in actuarial work. Naturally actuaries began to wonder whether such computers could be of assistance to them in developing insurance premiums, in mortality investigations and in other such activities. However, when an actuary tried to find out how to apply such equipment he quickly learned that life insurance people and electronic engineers were two groups who did not speak each other's language. He found the engineers quite willing and even anxious to have their ability and experience applied to the changes necessary to take computers out of purely laboratory work and into the business world, but they were lacking an adequate picture of the facilities the business world needs. On the other hand, actuaries and others in the life insurance field also lacked a sufficient understanding of the equipment which might reasonably be expected for insurance use. Some medium was necessary to bridge the gap between the two.

It would obviously be undesirable and unnecessary for each of us who might be able to use such equipment to start from scratch and separately begin to build up the knowledge needed. The preferable approach seemed to be through some joint effort. Accordingly, five years ago the Society of Actuaries appointed a committee[1] to examine into new recording means and computing devices and to report when it felt that such devices had been sufficiently developed that life insurance companies could consider their possible employment. Two such reports were made last year, one at the regular Spring Meeting of the Society of Actuaries in Washington, D. C., and another at a special meeting held in New York in September, 1952.

† Vice-President and Chief Actuary, Metropolitan Life Insurance Co., New York, N. Y.

[1] Committees were also appointed by the Life Office Management Association and by the Insurance Accounting and Statistical Association. Recently the British Institute of Actuaries also appointed a committee.

Studies by Committee of Society of Actuaries

The committee did a great deal of work in developing detailed procedures for applying electronic data processing systems to insurance operations. On the basis of the information acquired to date, it believes that systems suitable for day-to-day use are now available and that with such equipment substantial reductions in operating costs can be made. Further, it believes that the large potential of such devices will not appear clearly enough for practical purposes until some companies actually begin using them and make their experience generally available. A review of some of the many considerations and studies which produced these beliefs, and which led in at least one insurance company to a decision to contract for a magnetic tape system, may help to indicate the attitudes businessmen are likely to develop in considering the use of such equipment.

At the time of the appointment of the Committee of the Society of Actuaries there were no new systems actually available to business. Projected ideas which stemmed from laboratory use of computers like the Mark I, the Eniac and the Bell Relay Calculators, however, suggested quite strongly that corresponding equipment suitable for office use could be developed. The projected ideas were very good ones indeed—so good that even before the equipment became available, the committee tried to develop ways of using it.

First Approach

We started our studies by accepting, as a reasonable expectation, certain proposed plans for a magnetic tape computing system. With this visualized equipment as a starting point we tried to develop just what an average life insurance company might do with it to operate more efficiently. A very detailed operating plan for a hypothetical company was worked out on paper—but it did not supply a basis for any action. In effect, it said: "if a certain visualized magnetic tape system were brought into existence, it would make possible some radical changes in life insurance operation. These changes would greatly reduce the cost of current operations but it cannot be determined whether such changes can safely be made, how much they would save or when they would become possible." Needless to say with so many unknowns in the picture, there was little that could be done with this effort. Another approach was needed.

Second Approach

Several potential suppliers had been asking for a description of insurance practices. The thought was that, with a complete enough statement of current insurance practices, the suppliers would be able to design equipment especially suited to insurance work. Our committee felt that perhaps it might be useful to prepare a detailed statement of insurance procedures and make it generally available. This was tried. Much effort went into such an attempt but the project had to be abandoned because:

1. There is too great a difference between the practices, organizations and policies of the various companies to permit a description of general applicability.
2. It would be asking too much of an already burdened group of project engineers to require them to take on the formidable task of becoming sufficiently familiar with the insurance business.
3. It would be much easier for those of us who knew the insurance business and had a working familiarity with punched card techniques to develop the applications. The development of automatic procedures for the insurance business requires a deep insight into practices and objectives. It requires familiarity not only with the normal run of work but also with the variations and unusual circumstances which must be provided for.
4. Finally, it might be very misleading to supply indications of how work was now being done. There appeared to be good reasons for believing that substantial changes in current practices would be necessary to make effective use of the expected new equipment.

Third Approach

Out of this effort came the very strong notion that we have to avoid generalizations. Accordingly, it was decided that to be specific in indicating possibilities, we should confine ourselves to machinery already in existence and its possible effect in a particular company. This attitude led to a third approach which concentrated on the practices of one company. We first considered the manner in which a punched-card electronic computer could be used on certain work in order to establish a minimum standard for comparison. Then we visualized the manner in which a magnetic tape system might accommodate the same area of work. This effort resulted in development of a plan for administering life insurance policies suitable for use with several different electronic data-processing systems. This plan has become known as the Consolidated Functions Plan. A fairly complete description of it appears in publications of the Society of Actuaries.[2]

[2] In the Spring 1952 Transactions of the Society of Actuaries and in a report entitled "Report of Committee on New Recording Means and Computing Devices, September 1952." May be purchased from the Society of Actuaries.

Guide Posts

Perhaps the most important feature of this plan is the far-reaching change in current home office organizations and practices which it suggests. Principles behind this plan of general application may be stated as follows:

1. Apply the system to a whole job—not only to a departmentalized piece of it.
2. Combine small jobs with others to create the volume of work necessary to use the system efficiently.
3. Consolidate source records (i.e. combine files) so as to economize on the effort required to extract the data to be supplied to the system.
4. Perform all the work required from one handling of the data.
5. In devising the methods to be used, substitute arithmetical and logical operations for record look up, table look up and fact-recording operations as much as possible. Confine the information of record to the basic items. Avoid recording the associative ones which can be derived as needed.
6. Use a system in which sufficient checks have been included to avoid a cumbersome superstructure of clerical controls and error-correction routines.

It is to be expected that large economies in information handling would result from consolidating records and from combining work now done at different times and in different places. The plan emphasizes that such changes can be managed with equipment already in existence.

This, very briefly, covers the activities of the Society of Actuaries Committee. To date, at least three life insurance companies have arranged to acquire magnetic tape computers. One of them intends to use a magnetic tape system on some of its actuarial work. In addition, a fourth company is attempting to devise a comprehensive magnetic tape operating plan covering almost all phases of its policy service work. The idea is to test it on a small segment of the company's business and to extend it as the pilot operation becomes sufficiently effective.[3] Others are now investigating at the detail level, thus suggesting that some of them may be close to the acquisition of equipment of this kind.

Office Routines in the Life Insurance Field

There is no doubt that within the next few years there will be very substantial use of electronic devices in the insurance business. Accordingly, I believe it would be well to spend a few minutes discussing the various office routines required for insurance work.

Much of the work to be done in an insurance office arises out of the necessity to render service to policyholders. They are notified regularly of the premiums or benefits due. Upon request, adjustments and policy settlements are made as soon as possible. Many trans-

[3] A description of this plan appears in the Spring 1953 "Transactions of the Society of Actuaries," p. 198.

actions are handled each day by procedures which are very different in their details. In a broad sense, however, these procedures can be generalized by saying that each policy transaction to be dealt with involves some or all of the following steps:

1. The extraction of data from the information which initiates the transaction (a remittance, a letter, a notice of death, and so on).
2. The extraction of data from the information of record (as for example, reference to policy card files).
3. The use of extracted data to develop new information. (For example, putting data from (1) and (2) into a computer to determine a policy's cash value.)
4. The preparation of the documents required to execute the transaction. (The writing of checks, notices, statements, etc.)
5. The recording of the transaction in the recordkeeping and accounting system.

When a policyholder changes his address, only steps (1) and (5) are necessary but when a policyholder requests his policy value in cash, all five steps must be applied. On certain regularly repetitive activities such as sending premium notices, the transactions are initiated by the company. Data processing systems appear to offer their greatest potential in the automatic performance of the operations required in step (3). Steps (4) and (5) are aided somewhat by such systems, but very little gain, if any, appears to be possible in steps (1) and (2).

To apply such procedures, a company must maintain a record of many facts for each policy on its books. In order to do this with clerical organizations, separate files and a separation of functions into manageable pieces is a necessity. In one company this involves keeping at least ten separate policy files spread throughout several departments. There are separate files for billing purposes, separate files for dividend purposes, separate files for policy loans, separate files for actuarial purposes and so on. The maintenance and extraction of the information of record is currently a fairly sizeable task because of the large number of separate files in existence and the many references being made to such files. This suggests that a way of keeping the record which permits automatic extraction—some random access device—would be very valuable. This may be, but it must also be kept in mind that procedures involving many separate files are not necessary with modern data processing systems. When such work is mechanized with automatic equipment, consolidation becomes indicated. There have been studies suggesting that perhaps as much as 80 per cent of the number of file references can be avoided by merging separate files and changing office procedures to a basis better suited to electronic systems. If such a reduction in the number of file references can first be arranged by consolidation, then a random access device (an automatic file) would be required to operate on only 20 per cent or so of the original work load. On such a substantially reduced volume, the appeal of automatic extraction appears to become very limited in so far as possible economy is concerned.

The preparation of statements, checks and other such documents represents another area of insurance activity for which high speed devices are currently receiving a great deal of attention. In one company a count was made of the number of checks, notices, statements and so on which are currently being prepared. This suggested that, while the inclusion of high speed printing in a data system would increase the savings possible, the size of the increase was relatively small. The need for separate printers of speeds very much higher than present punched card tabulator speeds is still a debatable proposition, at least as far as the insurance business is concerned.

Reference has already been made to a Consolidated Functions plan. I will not attempt to describe it here other than to say that, as previously indicated, it suggests a large reduction in the number of card files being kept and a combination of much work often done separately in different departments.

This plan, operated by punched card equipment, indicates roughly the potential possible short of introducing tape processing equipment into insurance operations. In one company, if it were operating on such a plan, estimates indicate that the cost of servicing insurance policies would split roughly 20 per cent for extraction (items (1) and (2) of the generalized procedure), 65 per cent for processing the data and recording the results in the recordkeeping and bookkeeping system (items (3) and (5)) and only 15 per cent for preparing documents. This suggests that substitution of a tape system would yield by far the most economy in the processing area with correspondingly small gains to be expected in filing and printing operations.

Magnetic Tape Files

Perhaps the most radical idea which business is being asked to accept is the idea that a reel of tape can safely be used to carry information now being entrusted to visual card files. Some of us have already accepted the use of magnetic tape as a processing tool. We have also accepted its use as a substitute for secondary card files but we cannot quite agree that, at present, it can safely be used to carry irreplaceable information now being recorded on primary card files. The adequacy of tape for this purpose has not yet been sufficiently demonstrated.

We are not quite sure that the tapes now in use or being tested are sufficiently safe from accidental erasure, loss of information through breakage, kinks, dimensional instability, flaking, and other such occurrences. Nor have we been satisfied that the devices currently being employed to read and write on magnetic tape can be relied on to do so with accuracy. While the error rate is undoubtedly very small with such devices, the exposure to such errors is tremendously increased when it is realized that the maintenance of a magnetic tape file generally involves rewriting a whole reel of tape each time item changes are to be recorded.

In addition to this question of adequacy, there are some practical considerations to be dealt with. In the life insurance business, our practices are subject to review and supervision by the Insurance Departments of the various states. We have no way of knowing, for example, how long it would take for policy records in magnetic tape form to be acceptable to the state departments. Nor do we have any way of knowing how long it will be before the courts will consider acceptable, evidence existing in magnetic tape form. Microfilm evidence is still not acceptable in some courts. We think therefore that there still is some doubt as to the adequacy of tape records themselves—but even if this doubt is resolved, it will probably be a long time before the business community at large will consider tape records completely acceptable. For these reasons, we feel that insurance procedures intended for use in the next few years should not depend upon magnetic tape as a primary recordkeeping medium.

In developing the Consolidated Functions plan already referred to, a question arose as to whether we should suggest that the basic policy files be maintained in punched card form or in magnetic tape form. We chose to suggest punched card files—not only for the reasons already indicated—but also because of the relatively small increase in expected savings that could be attributed to the substitution of tape files for punched card files. We reasoned that the value of magnetic tape policy files over punched card policy files would be limited to not much more than the cost of transforming the card record to tape and vice versa, each time it was used. Under the consolidated plan such transformations are held to a minimum. To develop a measure of this, an estimate was made of the plan applied with a magnetic tape computing system and punched card policy files as compared with the same plan applied with a magnetic tape computing system and magnetic tape policy files. The estimated difference between these two ways of operating was small. The size of this difference among other considerations led to the opinion reported at the 1952 Eastern Spring Meeting of the Society of Actuaries to the effect that "at least so far as the life insurance industry is concerned, perhaps not more than 10 per cent of the entire potential possible with electronic devices can be attributed to an automatic file."

This analysis of punched card files versus magnetic tape files left the impression that tape files should not be counted on until extensive experience with tape as a processing tool—as a secondary record—has been acquired. With converters available to link existing punched card operations with tape processing systems, the need for introducing tape policy files does not appear to be a governing one. A very substantial amount of economy can be arranged through the use of tape computers—without the necessity of risking operation with tape files until more extensive demonstrations of their practicability have been made.

Limitations of Current Systems

In developing insurance procedures the choice of the methods to be applied is governed to a great extent by the capacity and limitations of the data processing system to be employed. The number of files to be kept, the manner in which they are to be maintained, the type of operations to be used and other such matters are practically prescribed by the necessity of avoiding the system's weaknesses while making extensive use of its strong points. For example, a sorting method which involves a minimum amount of writing on tape is to be preferred when the tape writing mechanisms are slow and less dependable than the internal components.

In developing the Consolidated Functions plan, it was necessary to take into account some of the limitations of the data systems contemplated. So far as the operation of the plan with current tape systems is concerned, the following devices were built in to overcome the more important of the system's limitations:

1. The available data systems do not include the facility of random access, except at prohibitively slow speeds. Accordingly, visual card files were retained to permit clerical references to be continued under the consolidated plan.
2. The available data systems do not lend themselves readily to the task of keeping a historical record of the account with policyholders. Accordingly, posted records produced by punched card tabulators were included in the plan.
3. Tape systems do not appear to handle sorting very economically, therefore, the procedures were designed with the idea of keeping the amount of sorting required to a minimum. One of the devices employed was the inclusion of a punched card policy file maintained in an order which would avoid sorting by the tape system. Such an arrangement, however, makes the file difficult to use for other purposes.
4. The available systems require regularized maintenance and testing to keep them in efficient operating order. Breakdowns, as with most machinery consisting of thousands of parts, involve difficult diagnostic work to analyze the difficulty and, sometimes, an extended repair job. Under the circumstances, it is questionable, at least, whether work which must be completed on an hourly or daily cycle can safely be entrusted to them. Thus, a monthly work cycle was built.
5. The available systems cost a great deal. For companies other than the very large one studied, a relatively small work load would be imposed by

the consolidated plan. This suggests the use of a service arrangement under which a computer might be rented for temporary periods. A number of checks and balances were built into the plan to permit safe operation with an outside service arrangement.

6. The available system transfers a great deal of control to a small number of persons associated with the operation of the equipment. Different approaches to internal auditing and checking may be indicated. More frequent checks on the accuracy of policy records are indicated. Periodic procedure reviews and test checks on performance and results achieved will also be necessary. In the Consolidated Functions plan, extensive consolidations are contemplated. However, three different departments, each with its separate records, are still retained. This is done primarily to include a basis for proper internal checks and balances.

This, of course, is by no means a complete list but it does serve to indicate some of the practical necessities to be kept constantly in mind. It is not solely an automatic processing problem that must be dealt with. It is a matter of devising a mechanization compatible with the current business climate and acceptable to the business community at large. Also, it must be one which can adequately be justified on economic grounds. Speed and availability of data are secondary considerations in the life insurance business.

Acquiring a System

A company which wishes to obtain first-hand operating experience with an electronic data processing system is faced with many questions. The most difficult to deal with is the matter of determining whether an electronic system should be acquired; and, if so, which one it should be. A prospective user of this equipment appears to have three choices. He may choose a system designed to operate from punched hole devices (punched cards or punched tape) thus limiting the speed and the degree of automation possible while keeping the price low. He may choose instead a system designed to operate from magnetic tape, thus reaping the advantages of greatly increased speed of operation. His other choice is to sit out the developments a little longer. This would be done with the expectation that future equipment will permit more automatic processes than now possible. Just which one of these should be followed depends on many practical situations within the individual companies. It also depends on whether the use of a system requires outright purchase, full-time rental, or an hourly service charge from a computing center. I believe you will be interested in the reasoning which led one company to contract for a magnetic tape system.

A Punched-Hole or Magnetic-Tape System?

This company operates with procedures that are mechanized to a fairly high degree and is currently engaged in several programs for converting additional clerical activities to a punched card basis. To settle on one of the three choices mentioned above, it was assumed that the Consolidated Functions approach supplied a rough measure of the degree of automation to be ultimately expected of data processing systems. Rough estimates were made of the cost of operating the plan after it had been completely installed and was operating on a company-wide basis. The estimates were made for operation with a punched card computer and punched card files, and for operation with a magnetic tape computer and punched card files. A comparison of costs favored the tape processing system but the difference was small. A choice could not be made on cost alone.

These studies made the consolidation idea appear very attractive, not only as a means of reducing operating costs but also as a means of streamlining the organization along workable lines. However, on the other side of the ledger, there were a number of deterring factors. Such a program would take a long time to install. The gains would appear only after a fairly extensive conversion had been made. The conversion would require producing punched card records for several million life insurance policies now recorded in typewritten card form. The loss of a very substantial investment in an existing recordkeeping system had to be taken into account. In addition, the capacity of the company to make the required changes was being taxed to a considerable extent by the several conversions to punched card systems currently being made. To these deterrents must be added the element of risk which exists in introducing equipment as yet untried on any commercial application. The company, therefore, although inclined to accept the idea that some sort of consolidated operation was indicated as a long-range objective, still felt that a system should be applied to a localized area as a means of getting started.

The company has a large area of actuarial work which is currently being done by an assembly of more than 100 separate punched card machines (involving a yearly rental of about $225,000). This work involves the development of the insurance statistics needed for the company's financial statement and for various experience analyses. This area of work is of such a nature that it permits the introduction of an electronic processing system with a minimum of disturbance and risk. It constitutes a severe test of the tape system because it involves a very large amount of sorting, an operation on which tape systems are admittedly expensive. It was found that a data system designed to operate from punched hole designations does not offer material gains over regular punched card equipment on this particular work. Studies and tests indicated, however, that a magnetic tape data processing system would offer substantial economy on this actuarial work. In addition, the magnetic tape system offers the added capacity needed to cope with certain non-repetitive types of special work such as is involved in determining new premium rates or in preparing new dividend schedules.

In this connection, it should be emphasized that the company involved had already consolidated its actuarial work for several different lines of business. Such an assembly of punched card equipment applied to producing insurance statistics would not be found in other companies.

Additional considerations favoring a magnetic tape system were found in the fact that such a system would be operated by a small number of people and would require fewer separate machines and less movement of work; also, in the expectation that fewer mechanical errors would occur with a system containing built-in checking devices.

Primarily, however, this company chose a magnetic tape system over a punched hole installation for one extremely important reason—the system could be gainfully introduced into this particular area without requiring a major reorganization elsewhere as a condition precedent. This provided the facility for a first-hand understanding of its capabilities and limitations through actual use, without risking the entire recordkeeping system of the company on its success. The fact that it could also be used in connection with a plan like the Consolidated Functions approach, the fact that it would be operated by a small staff and supplied desired reserve capacity of course weighed very favorably in the decision.

An Existing or Future System?

Having made this choice, it was still necessary to consider whether to acquire an existing tape system or wait for some future system which might be less costly or more versatile. An attempt was made to visualize in what respects a future system might surpass the existing ones. It did not seem unreasonable to expect that future systems would be more completely automatic. For instance, it was assumed that, instead of extracting information from files as a separate operation, devices for the automatic extraction and processing of information would probably be built in. It was likewise assumed that instead of getting answers first in tape form for separate printing on a different machine, a printing device tied into the system would be available. On the basis of this degree of automation an attempt was made to estimate, very crudely of course, how much more savings might occur if the Consolidated Functions plan were in operation; that is, the cost of operating this plan with an existing magnetic tape system was compared with an estimate prepared to reflect operation with the more completely automatic tape system assumed for the future. For lack of any better indication, it was assumed that an automatic tape system with built-in files and built-in printing mechanisms would sell for no more than an existing system without such built-in devices. This comparison suggested that perhaps 90 per cent plus of the total potential possible with such highly automatic future devices could be realized by the lesser degree of automation already possible with existing equipment.

Of course such a suggestion immediately leads one to wonder a bit. Perhaps a disproportionate amount of effort is being expended in a desire to make systems completely automatic. Maybe some of this effort should be directed toward pilot applications which will supply some field experience to work with. To us in the insurance field, the development of procedures to accommodate semi-automatic systems appears more certain than the development of more completely automatic data systems at a practical cost within the next few years.

The other consideration which entered into this question of "buy now or wait?" was an estimate of how long it would be before the company would be able to recapture the cost of a present system out of expected savings. After allowing for the necessity of starting out on a local operation, and the additional loss to operate on both the old and new basis at the outset, this period was estimated at somewhat less than four years. With such a rapid rate of buy back, coupled with the large proportion of the total realizable with existing systems, the question was resolved in favor of the introduction of an existing tape system.

This covers some of the attitudes developed by a group which is moving from the investigating phase into actual operation. Unfortunately, the only available basis for decision involves future projections and conjectures. With some life insurance companies acquiring tape computers within the next year, this projection basis may soon be displaced by actual experience. This will be valuable because visualizations have done about as much as they can. Actual use of current models on day-to-day work appears to be essential to the development of both the proper applications and the improved processing systems of the future.

Indications for the Future

It has been difficult for a potential user of a data processing system to decide what to do. He realizes that the greatest impetus to the development of improved systems will stem from day-to-day operating experience but he also is conscious of the possibility that any system acquired now may be out-of-date a few years from now. His plight is not much relieved by the many projections being made in engineering circles—promising very complete automation; promising transistorized computing systems; promising inexpensive small units; and so on.

Those of us who must deal with this matter need some good indications as to what and when you engineers believe you can deliver. Let a picture of a transistorized computer appear in one of the popular magazines and we immediately have to deal with the question of whether or not the present vacuum-tube computers are obsolete. Let some ideas as to photographic storage appear on the horizon—and we begin to wonder whether they are one, ten or twenty years away. It is in this area that we need the most guidance and it is the one in which we will probably receive the least in the form of concrete information.

From the standpoint of the life insurance companies, the guideposts which have evolved out of our studies imply a basic reengineering of present procedures. Current organizations, methods and attitudes will no doubt have to be reconsidered, both in detail as well as in broad perspective. It is hardly likely that a company would embark on such an undertaking without acquiring some first-hand experience in operating a data processing system. Accordingly, the most one can expect from paper studies of possibilities, no matter how carefully they are done, is a general goal toward which a company might move.

As already indicated, some of us will soon be performing our day-to-day work with highly powerful data processing systems. We are currently engaged in trying to plan for the future. We have begun to view as more urgent:

1. The importance of identifying the basic principles and objectives of our business and then thinking through, right from scratch, the practices that have evolved. With radically different tools, entirely different practices may be needed.
2. The acquisition of the skills and knowledge required to organize satisfactory procedures for using these important new systems. It is clear that a user cannot order the equipment, plug it in and hope to have an automatic office system. It is also clear that he can no longer expect the benefits of clerical operation. Clerks are capable of interpreting his intentions and judging the results achieved all along the way. Machines will not do this; therefore he must think through every step completely and prescribe the procedures in minute detail.
3. The necessity of reducing the amount of planning effort necessary to develop the procedures required by electronic systems. This might well involve the creation of a library of worked out routine (on an inter-industry basis if possible). It also suggests that the procedures be developed not only for the particular job on hand—but also for the general class of work to which the particular job belongs. Perhaps suppliers should furnish a kit of standardized computer routines as part of the equipment package.

A sizeable task lies before us and in it we must be constantly on guard not to project current rationalizations too far ahead. There is much yet to be discovered about operation with equipment of this kind.

Conclusion

A very remarkable amount of progress has been made in the 10 years or so since the laboratory prototypes of current data processing systems were developed. Some of us in the applications end of the picture have come to the point where we believe that about as much as can be accomplished by paper planning has been done. We are ready to apply the new tools to day-to-day work. We are conscious, however, of the sizeable undertaking which is involved to recast our current operations into the form best suited to highly automatic equipment; therefore, we expect a relatively slow gradual process of accommodating business needs. In the life insurance business, we are just beginning to see a reasonable amount of effort being directed into possible uses of these new and radical tools. Some of us believe that by far the major portion of the ultimate potential can be achieved with systems that already exist. We are therefore inclined to the view expressed by Professor P. M. S. Blackett[4] when he said:

> " . . . relatively too much scientific effort has been expended hitherto in the production of new devices and too little in the proper use of what we have got. . . ."

Parallel with your development of new and improved components for data processing systems, extensive investigations into the manner in which they can be used should be conducted by potential users. They must, of course, take into consideration the necessary gradual manner in which radical changes should be introduced. They must also emphasize strongly the fact that business decisions are based on economic tests with adequate consideration for the human elements involved.

[4] "Operational Research," paper published by British Association for the Advancement of Science, vol. V, no. 17; April, 1948.

Discussion

R. F. Osborn (General Electric Co.): Have groups similar to the Actuarial Society Committee been organized to make studies of general business applications? If not, shouldn't one of the conference sponsors arrange for such a study?

Mr. Finelli: The Life Office Management Association and the Insurance Accounting and Statistical Association, two associations connected with the insurance business, have committees which are making studies. There are corresponding committees in existence for the electric light and power companies. Whether another committee should be appointed to cover the same general ground is a question for you to determine.

E. H. Friend (U. S. Navy): What plans are made under the consolidated plan for the inevitable breakdown period?

Mr. Finelli: Two devices are employed. First, any work which requires daily or hourly service is not made dependent on closely scheduled computer performance. For example, to accommodate those people who wish the cash value of their policy, the plan calls for computing this value in advance on all policies. This cash value is recorded on a card record. When a person applies for it, reference would be made to his card record, a few simple adjustments made, and then the payment would be arranged. No computer operation intervenes between the time he asks for the money and the time it is paid. Second, the plan has been arranged on a monthly work cycle with a computer load requiring less than 50 per cent of the available computer time. In general, it reflects the idea that computers can safely be relied upon to produce a month's work within one month's elapsed time when the work load does not exceed 50 per cent of the available computer time.

R. D. Dotts (Pacific Mutual Life Insurance Company): Was the 90 per cent figure based on a full consolidated functions approach (not actuarial alone) and on present equipment without a high-speed printer?

Mr. Finelli: The answer to both parts is yes. It is based on the full plan, not only the actuarial area. The answer is based on the use of current punched-card equipment without including the savings which might be achieved with a high-speed printer. Of course, it must be kept in mind that current punched card tabulators are speedy printing machines, so fairly high printing speeds are reflected in the 90 per cent.

D. Mitchell (Bell Telephone Laboratories): Would it be worthwhile to be able to transmit data over distances up to hundreds or thousands of miles?

Mr. Finelli: Sounds good but I don't know. One life insurance company operates on a somewhat centralized basis. Most of its records and processing are done in New York. Some duplicate records are kept in District Offices. Just how records might be eliminated or costs otherwise cut in such a company by long distance communication facilities is hard to see, offhand. Another company, however, which is more decentralized and which has several scattered offices with full recordkeeping and servicing activities might find such transmission very useful indeed, particularly so if it can feed into a central computing center. The question would have to be investigated very thoroughly with due attention to economic factors, before an answer can be developed.

W. B. Hebenstreit (Hughes Research and Development Laboratories): "Random access" is not precisely defined. What do you have in mind with respect to (1) maximum time of access to a single account? (2) average rate of access? (3) size of files?

Mr. Finelli: On (1), access to an individual policyholder's account within a minute or two minutes is, in general, the kind of random access I have in mind. This, however, should not be regarded as a rigid requirement in that operating systems involving a longer wait time can be devised.

I am not sure what you mean by (2), so I will have to skip that.

As to (3), the size of the policy files required for the largest life insurance company can be judged by the fact that there are some 5 million plus policies in the Ordinary branch—that is, the business that's billed for (payable annually, semiannually, etc.). That would involve, figuring roughly 300 characters per policy, $1\frac{1}{2}$ billion alphameric digits. Now that's only one of four branches of business. I would say that if you multiply that $1\frac{1}{2}$ billion by a factor of 5, you would get a picture of the over-all record-keeping load in that company.

S. B. Williams (Consulting Engineer): Has consideration been given to the use of some form of microfilm for the primary files?

Mr. Finelli: Not too much. We've heard references to the possibilities of photographic techniques. I think there is some work on the coast along those lines but I'm not sure how far it has developed.

S. Meyer (Remington Rand): In regard to the necessity for keeping primary records on proven nondestructible forms, could you comment on the use of a magnetic-tape preparation machine which produces a typewritten copy as a by-product?

Mr. Finelli: If the machine is low priced, it is a step in the right direction because it seems necessary to bring the cost of direct recording on magnetic tape down to a figure more competitive with other recording means. The fact that such a machine might also produce typewritten copy does not affect very much the question of adequacy of tape in the business sense. Not all the information placed on tape would be recorded by the typewriter—some would be placed on as part of a calculation by the computer. Further, it would pose an extremely difficult operating problem to try to support information recorded on tape by a paper typewritten record of the same facts. The tape difficulties we are not sure of yet are things like flaking, dirt accumulation, accidental erasure by an operator, kinking due to either operator or machine handling. We do not known that these are insurmountable problems, but we need a great deal more information and experience before we can be sure that they are not.

J. E. Mekota (Raytheon Manufacturing Company): Where may copies be obtained of the Actuarial Society Publications?

Mr. Finelli: From the Society of Actuaries offices in Chicago.

Mr. Mekota: What proportion of present sorting operations can be eliminated according to actuarial studies?

Mr. Finelli: I don't have a proportion to give on that but there is a great deal of it—a very large proportion that can be eliminated. There are two general devices—one, of course, is the one that was indicated earlier: arrange your files in such a way that for most of your statistisal work it becomes a matter of counting up the files just the way they are. The second one is to adopt the device of dispersion and collection— in other words, when you have a computer that can count and add up say in a thousand separate pockets, you don't need to classify to the degree necessary to add up only in one pocket and deal with it serially. You could kind of spread, let us say, a count of your insurance outstanding for each age over 60 or 70 pockets. This would avoid the pre-sorting that would be necessary to arrange the data first by ages, and then add it, one age at a time, through one pocket. I think a very high proportion of sorting can be eliminated by proper design of procedures —but I can't give an estimate of it at the present time.

Computer Applications in Air Traffic Control

VERNON I. WEIHE†

WE AMERICANS ARE a capricious people. We warm up and develop tremendous enthusiasm for an idea, an activity, or a particular way of doing things. We expect a great deal from the objects of our enthusiasm, and when our great expectations fail to materialize on schedule, we cool as rapidly as we once warmed.

Six years ago the RTCA SC-31 "guide plan" was published after eight months of enthusiastic effort by some of the better operational and electronic-engineering men in aviation. This guide plan envisaged the attainment of a fully automatic air-traffic-control system using analog and digital computer techniques, to become completely operational by 1963. This phase of the program has dropped behind schedule and many of us—because of our capricious nature—have grown cool to the idea.

The need for automatic computation and automatic data handling is *immediate* and *urgent*. More urgent by far than we could have predicted six years ago. Our progress in the sciences of computation has been considerable; we are far more able to cope with the problems than at any previous time. Speed differentials between aircraft have spread beyond the limits of our most radical predictions. Rates of closing in aircraft

† Melpar, Inc., Alexandria, Va.

operations today have already outdistanced human capability using today's manual methods. Each month the requirements go further beyond the performance of our manual control system.

During the past two years the DeHaviland turbo-jet Comet has been placed in 600-mile an hour, routine scheduled Airline service. Other high-speed turbine driven aircraft will soon follow. Last month Crossfield flew a Douglas Skyrocket at 1,327 miles per hour in level flight. An American Airlines DC-7 made a nonstop flight from Los Angeles to New York at 400 miles per hour carrying passengers. During this same month, as if to put a bottom on the speed bracket, Piesecki unveiled the 40-passenger YH-16 helicopter—the type of helicopter which may well be the DC-3 workhorse of scheduled helicopter operations.

Commercial aircraft speeds today in scheduled operations exceed those of military aircraft of equivalent size and weight of World War II. Tomorrow's commercial aircraft will probably equal the speeds of today's military aircraft. The DC-7 now going into U.S. domestic scheduled service will permit a passenger to have lunch in New York and dinner in Los Angeles. All that is needed is more thrust and the speed will equal that of the Sun. We can have lunch in New York and arrive in Los Angeles in time to have another lunch the same day.

With closing rates now in excess of 10 miles per minute and rapidly approaching 20 miles per minute and higher, it is easy to see that manual methods of control using telephone, teletype, voice radio, and handwritten records are already outmoded.

Our present air-traffic-control system is operated primarily during instrument-flight weather. Only at terminals of high traffic density is there full-time control activity. But what about good visibility operations? How much longer can we rely upon "see and be seen" procedures to assure safe separation of aircraft under good visibility conditions?

Since the war there have been 172 air-to-air collisions during clear weather conditions. To minimize this type of hazard we should go to a 100 per cent, all-weather around-the-clock control system at the earliest practicable time. Our present manual system cannot cope with these requirements. Today's system is used for instrument-flight-rule traffic only. Economic penalties on both the user and the traffic control agency will not permit 100 per cent control using these manual methods.

The introduction of automatic methods into air-traffic-control centers and towers and into the aircraft cockpit will not be easy. There is too much prior art—all of it manual. The superior features of mechanical brains for certain air traffic control functions will be hard to prove to the personnel most directly involved. Their comparatively unfavorable attitude is easy to understand. Especially considering the fifteen-year history of essentially safe instrument-flight-rule air-traffic control, and recognizing unacceptable operational concepts contained in several of the automatic air-traffic-control papers and reports published since 1948.

Although this unfavorable condition is readily understandable, the consequences of it are nonetheless serious. In its most dramatic form, if the occasion should arise, it could contribute materially to successful attack on the United States. In its least dramatic form it could merely reduce the safety and efficiency of peacetime flight and have a controlling influence on aircraft-flight economics and the types of aircraft chosen for various types of service.

What is our present system like?—and what are some of the aspects which support arguments for the introduction of automatic system elements?

Tower and center personnel communicate with each other by voice and teletype. They communicate with pilots using voice radio. They use procedures described in the ANC manual with local variations as deemed advisable.

Records of flight progress are kept current by writing with pencil or posting strips; records for pre-flight and postflight use are made by teleprinter and magnetic-tape recorder. Flexibility is the keynote. Empirical trial-and-error experimentation is the order of the day. Quick fix after quick fix has produced a system which has grown like Topsy. The very flexibility of expression possible in the system makes accurate detailed description of it impossible. Quick arithmetical calculations plus experience based upon repetition provide the means whereby the air traffic controller checks on the aircraft's anticipated progress and determines whether or not conflicts in separation are imminent.

Prior to issuing the initial clearance which permits an aircraft to enter the system, the traffic controller is given a detailed written record containing such things as point of departure, desired destination, desired routing, aircraft identification (type, serial number, speed, owner, and so forth), fuel range, and approximate time of take-off or entry into the control system.

After the flight is airborne the pilot reports altitude over prescribed compulsory navigational fixes. His progress and ground speed are thereby determined. Any special situations or conditions which arise are communicated between controller and pilot via voice radio.

Pencilled notes are placed on posting strips which are moved from top to bottom on the posting board as the aircraft proceeds into and through areas of cognizance of the various controllers.

Safe longitudinal, lateral and vertical separation, and expeditious sequencing and flow are sought through repeated scanning by the air traffic controller of posting strips. After deciding what clearances to issue the controller communicates—directly or through an intermediate communicator—with the pilot. Transfer of control jurisdiction and other pertinent information is communicated between controllers using voice and teletype facilities. As the faster aircraft tend to approach the speeds involved in manually handling their progress messages, the time remaining for co-ordination between controllers in dense traffic areas all but disappears. The co-ordination problem grows more and more complex while the time for accomplishing co-ordination is being slowly whittled away.

What steps can be taken to overcome the unfavorable attitude and to plan a program leading toward automatic air-traffic-control computation and message handling?

First, we have the difficult problem of establishing what basic principles and practices of air traffic control we will utilize. The most thorough studies to date have led to system proposals containing elements which were unsound operationally, technically, and economically. Since we cannot talk to computers and have them talk back to us using our present voice facilities, we need a specialized point-to-point and air-ground (ground-to-air and air-to-ground) communication system. We also must feed in aircraft position information, aircraft identification and communications routing information. Answers from the computing devices should be flashed directly to the interested pilots and controllers.

The problem of applying automatic computation and data handling to air traffic control is so complex that many technical approaches are applicable. No doubt there are also a number of acceptable solutions.

However, some of the proposed solutions are obviously unacceptable. Analysis of certain aspects of these proposed solutions provides some guidance for the future.

In its basic concept an acceptable plan should contain a proper balance between requirements for communication services and requirements for computation. Human and economic factors also must be properly weighed.

A recognized engineering group conducted a series of studies and published a group of reports covering the problems of automatic air traffic control. Several basic conclusions contained in their proposals were unacceptable operationally and unsound economically.

For example, it was proposed in one of the published reports that aircraft be admitted into the system only after an increment of runway time at the terminal of intended landing has been assigned. When low visibility and low ceiling conditions were forecasted for the arrival period—the right to take off was to be denied.

Traffic congestion was to be eliminated by restricting takeoffs. This simplication of the problem does not satisfy operators of aircraft. They desire to prevent traffic congestion by moving traffic more expeditiously —not by further restricting its movement. They also do not wish to predicate operations upon the validity of local area or so-called micrometric weather forecasts.

In this same report the computer-programming concept and the communication-service requirements led to an economically unsound end result. About a dozen major centers were to be implemented at various points in the continental United States. Each center was to be connected to all of the others and to the many ground aids and airports of the system using wide-band (microwave or coaxial cable) communication links.

The plan called for the initial aircraft clearance to contain the flight track, and time of use through each computer control area. Moment by moment as the flight progressed the computers were to probe through the complete space-time pattern for each flight to establish whether or not potential conflicts existed. If a potential collision hazard was indicated, the computer was to issue appropriate changes in the flight plans of the aircraft involved.

Although this plan is operationally unsound by reason of the fact that changes in flight plan would be made at unnecessarily frequent intervals, its main flaw is the economic burden of the elaborate communication system needed. An excessively large proportion of the communications would serve no useful purpose. Communications bandwidth requirements in excess of minimum needs should for economic and several other reasons be avoided.

Variations in flight progress due to aircraft power settings, upper wind conditions and many other factors make the establishment of or adherence to strait-jacket time schedules a weak element on which to base a system. Flight economics are also adversely affected.

After the SC-31 Guide Plan was published, I attempted to block diagram an automatic air-traffic-control system to meet the operational criteria stated therein. During this effort I learned that human engineering problems will affect the system design in a variety of important ways. Suppose that we plan to build an automatic air-traffic-control system which will handle the peak of all peaks of traffic between two centers of high population density for some future time, say twenty or thirty years. Suppose also that we desire to utilize all operationally useful segments of airspace between these two points.

We will find that the system is overdesigned and that the air traffic controller is inherently incapable, because of basic human limitations, of clearing aircraft to efficiently employ an appreciable portion of the system's capacity. Codification of routes becomes useless when the number of codes gets too great.

Consequently, it appears essential that the number of aircraft, the speed of handling, and the number of flight paths at any one time be conservatively forecast. To these estimates a safety factor should be added, and equipment planned and installed to meet the requirements. Human factors and cost elements then appear to dictate the use of a building-block approach to the problem. Our philosophy of system design should not visualize the introduction of a plant the size of the New York City telephone exchange at a place the size of Altoona. Our system should be capable of growth with the problem at any particular location in the overall system. A system proposal which visualizes the general application of identical equipments and plant size at all locations will be difficult to justify. We should adopt a building-block philosophy.

Another serious problem concerns the question, how much information shall we display to the air traffic controller? And in what form? During the introductory period when there is a high ratio of manual to automatic handling a great deal of information will need to be displayed. As the ratio improves, less and less information will need to be displayed. More use of automatic facilities, and greater confidence will decrease the amount of information which will need to be displayed. The period of time during which extensive use of displays is made will depend to a large extent upon the reliability of performance of the actual equipments involved. In a safety service such as this, 100 per cent system reliability is of paramount importance.

Hence the computer and storage elements of the automatic system will need to contain features which permit quick and easy access to information for display. This especially applies to information related to the safe separation of aircraft. A completely automatic system which overlooks this point may be excellent, but without these evolutionary characteristics, we cannot get to it from our present manual system.

The above examples indicate that computer engineers will need to analyze practical operational problems in sufficient detail to outline and shape certain basic criteria to fit the technical aspects of the problem. Otherwise the technical problems will become prohibitively difficult. The computer engineer, then, will need to concern himself with air-traffic-control philosophies, procedures, and practices. He will need to have available quantitative information on airspace size allocated per aircraft, the amount and nature of the communication services available, and required, standardized identification and routing methods, codes, and so forth.

From this information he will need to determine the general sizes and locations of computer elements needed, and which elements of the problem are served best by analog techniques and which are best served by digital techniques.

To summarize broadly: we must accomplish our results in a manner which is compatible with current operations and will permit us to evolve toward fully automatic control as ground and air elements of the system are progressively installed. We cannot tolerate a switch-over date and must not decrease the safety and efficiency of operations during the period of transition. Nor must we increase the burden of the already overloaded controller during this period of transition. From today's viewpoint it also appears that we must accomplish all this in the face of a diminishing budget.

In addition to this operational plan—devised to suit the computation and automatic data-handling system—we also need more enlightened program support at policy level. It does not appear evident that policy makers are cognizant of the threat to America which may exist in our passive attitude toward automatic air traffic control. Twelve years ago at Pearl Harbor the lack of positive identification of friendly aircraft contributed materially to the Japanese success. Until the bombs fell, those in Control believed the aircraft to be a friendly flight from the mainland.

Since December 7, 1941 the basic procedures and tools of air traffic control have undergone little change. The small improvements which have been made can in no way match the progress made in the development of the aircraft, its bombs, and its weapons. Whereas in 1941 we could let a squadron or two of unidentified aircraft get through, today the lack of identification of a single H-bomb carrying aircraft could contribute to the destruction of an entire city.

Many of us who have been in air-traffic-control system planning for a number of years believe in and advocate the so-called common (civil-military) system. We believe that if we are forced to fight a war, we will make primary use of the operationally tried and proved communication, navigation, and traffic-control-system elements which are in use at the time the war starts.

We believe that all aircraft must eventually be positively identified—and be under positive automatic control at all times, regardless of weather or visibility. We believe this can contribute materially toward increased safety and efficiency of peacetime flying, while aiding those whose task it is to identify intruder aircraft.

Where do we stand at present in our efforts to adapt automatic methods to air traffic control? The Air Navigation Development Board and the CAA Technical Development and Evaluation Center have procured and are preparing to test some elements of the system. The Engineering Research Associates magnetic-drum unit is already at Indianapolis, and the Union Switch-and-Signal Indicator and Display system is scheduled to be sent there in a few weeks. Meanwhile the Bell System is making notable progress along with CAA at Indianapolis in the adaptation of teletype terminal switching gear to the *automatic data-handling problem*.

Within recent weeks there has been a re-awakening of interest in Washington. People important in these matters are again beginning to warm up to this subject. I believe 1954 will be marked by another serious attempt to get a re-vitalized and well integrated program under way. The techniques are available—we must formulate the operational problem more concisely, so that practical system elements can be assembled and programmed for installation. With the SC-31 "guide plan" as a starting point, I think we can, with the benefit of our new information, prepare a more realistic program; one which will have greater appeal to conservative operational people and can be accomplished with better economics in a reasonable period of time.

Programs for the improvement of air traffic in both technical development and in everyday operating services have shown slow progress in 1953. In many instances serious retrograde steps have been taken.

I cannot believe that a problem of such great importance to the defense and welfare of our nation can continue to be so badly mishandled by the clear-thinking people we have at the top aviation policy level.

In our capricious way we may very well generate a new enthusiasm, and replace the low performance of 1953 with an outstanding performance in 1954.

Discussion

J. E. Levy (Technical Consultant): Does the history of air collisions indicate the cause as inadequate flight programming or inaccuracy of airborne instruments?

Mr. Weihe: The 172 air collisions of which I spoke, took place under conditions of good visibility when no traffic control action was being exercised. Under such conditions it is each pilot's responsibility to maintain visual alertness, and to see and avoid other aircraft on collision courses.

M. Stateman (Sylvania Electric), **T. Brady** (Ford Instrument Co.): Where can I get a copy of SC-31 Guide Plan?

Mr. Weihe: The SC-31 Guide Plan was published by the Radio Technical Commission for Aeronautics, 1724 F Street, Washington, D. C. The price is one dollar per copy.

W. P. Byrnes (Teletype Corporation): At what speed per minute should a telegraph system operate in order to have it fit in with future traffic control systems?

Mr. Weihe: We have a group of different jobs to do in air traffic control. We have the initial clearance, programming of the aircraft through the system, and we have the problems of safe separation, and expeditiousness of flow. The teletype or telegraph requirements are different in each category. But I would say that we will unquestionably start with the speeds which we now have and at a later date certain parts of the problem may require speeds up to ten times the present speed. But meanwhile we have to cope with the speed problem by having more circuits; when the techniques develop so that we can use fewer circuits at a higher rate of speed then I think we will evolve towards the higher rate system. I would suggest that you talk to Mr. Gross of CAA TDEC on this subject; he knows far more about it than I do.

Data Processing Requirements for the Purposes of Numerical Weather Prediction

JOSEPH SMAGORINSKY†

Summary—**The physical background and historical development of numerical weather prediction is summarized. The nature of the mathematical problem and the work-load for a digital computer are then described. The existing system of processing meteorological information (particularly upper air data) for operational weather prediction is examined in the light of the anticipated needs tor operational numerical weather prediction. The critique exposes deficiencies which may be intolerable for operational numerical weather prediction. Some suggestions are offered for revamping the existing system utilizing modern technological advances.**

Numerical Weather Prediction

Historical and Physical Background

The central theme of this paper is the class of problems brought to light by the introduction of numerical methods for the purpose of weather prediction. What I will have to say must, out of necessity, be based on anticipated difficulties, since experience in *operational* numerical weather prediction by means of high-speed computers is nonexistent at this time. One can only speculate on the basis of experimental experience in numerical weather prediction and operational experience with conventional methods of prediction.

In preface, it will aid our perspective to give some background regarding numerical weather prediction itself.

The prediction of changes in the large-scale weather elements by means of physical laws has occupied theoretical meteorologists for many years. Lack of observations, an incomplete understanding of the physical laws and inadequate computational means prevented a break through in this problem until 1947. It was then that Rossby's basic work[1] of the late 30's stimulated Charney to develop a rationale for the prediction of large-scale atmospheric motion.[2] Only the availability of adequate synoptic observations and of high-speed computing machinery enabled him to carry his investigations beyond the initial stages.

The primitive hydrodynamic equations of motion for a compressible fluid in a rotating system reflect the fact that the atmosphere is capable of sustaining a wide spectrum of disturbances. However, for the purposes of large-scale short-range weather prediction (1 to 3 days), attention is focused only on those disturbances of planetary dimensions with periods of 3 to 7 days. For this scale of motions, the atmosphere behaves quasi-hydrostatically. Furthermore, because of the earth's rotation, the accelerations are relatively small so that the pressure gradient forces are approximately balanced by the Coriolis force. This "quasi-geostrophic" property of large-scale atmospheric motions together with the quasi-hydrostatic approximation when applied to the equations of motion have the effect of filtering out the meteorological-noise motions, that is gravity and sound waves.[3] The prediction problem is thus greatly simplified from both the observational and computational point of view, with scarcely any loss of applicability. For initial conditions, one needs in general only a knowledge of the three-dimensional mass (or pressure) field; it is not necessary to specify the wind field in addition, however, the geostrophic approximation permits one to utilize the wind field as an aid in determining the initial mass distribution.

All models thus far used for prediction purposes assume that the atmosphere is a thermodynamically

† U. S. Weather Bureau, Washington, D. C.

[1] C. G. Rossby, "Relation between variations in the intensity of the zonal circulation of the atmosphere and the displacement of the semi-permanent centers of action," *Jour. Marine Res.*, vol. 2, pp. 38–55; June, 1939.

[2] J. G. Charney, "The dynamics of long waves in a baroclinic westerly current," *Jour. Met.*, vol. 4, pp. 135–162; October, 1947.

[3] J. G. Charney, "On the scale of atmospheric motions," *Geofysiske Publikasjoner* (Oslo), vol. 17, no. 2, pp. 1–17; 1948.

closed system which slips along its lower boundary. First experiments in numerical weather prediction involved highly simplified models of the atmosphere. One finds that the flow at 20,000 feet or 500 mb closely approximates the mean flow obtained by integrating through the entire depth of the atmosphere. If one wishes to predict the 500 mb flow using information just at this level for initial conditions, it is necessary that the flow at all other levels be implied. To do this, we assume that the wind is parallel at all levels. This tacitly assumes that the lines of constant temperature in an isobaric surface are parallel with the streamlines. Thus, we have a model of the atmosphere which has only one degree of freedom in the vertical—normally referred to as the "equivalent barotropic model."[4] This may seem to be a gross oversimplification of the atmosphere which has no semblance to reality. In fact if we examine the constraint that is imposed on energy transformations we find that such a model is only capable of redistributing kinetic energy by means of dispersive processes—it is not possible in this model for potential energy to be converted to kinetic energy. As it turns out, this is not catastrophic. Large conversions of potential to kinetic energy occur only sporadically in time and in space in the atmosphere. They occur at the beginning stages of storm development, which may lasting for 12 or 24 hours, and in extreme cases for 36 hours. Developments of this type occur somewhere over the United States on the average of once every 3 or 4 days. Between times the flow behaves essentially barotropically.

The very first forecasts not only assumed barotropy but also assumed the disturbances at 500 mb to be small perturbations on a uniform constant west-to-east current and that these disturbances had a fixed character in the north-south direction so that the flow was essentially one-dimensional.[5] So simple was this model that predictions could be made by means of a desk calculator. Surprisingly, the results had sufficient resemblance to reality to indicate that a nonlinear two-dimensional barotropic model would give much better results. Such a model out of necessity required high-speed computing machinery to carry out 24-hour predictions. For this purpose, the ENIAC was made available in early 1950.[6] Under ideal operating conditions, it took 24 hours for a 24-hour prediction to be calculated for an area comparable to twice the size of North America. These tests bore out supposition that atmosphere tends to behave barotropically except for short but important periods during which storms develop.

It was thus apparent that if one expected to predict development, it was necessary that the atmospheric model be capable of making available potential energy which could be converted to kinetic energy. To do this, it is necessary to specify at least two independent pieces of information in each vertical aside from the boundary conditions. This is equivalent to spanning the vertical dimension of the atmosphere with at least two internal mesh points. Predictions with a two-level model were performed last year on the IAS computer.[7] The weather situation chosen for this series of calculations was the famous Thanksgiving Day storm of 1950 in the eastern United States. This storm was one of the most rapid and intense developments ever to have been adequately recorded by a modern observational network. The two-level model definitely improved the predicted motion, but the degrees of freedom of at least a three-level model were necessary to predict the full development which was observed to occur.

The two-layer results indicated that the accuracy of prediction by means of the simpler barotropic model degenerated much more rapidly as the forecast period was extended. One could then conclude that more realistic models would probably yield good results for forecast periods of 36 or possibly even 48 hours.

A three-layer model embodying quasi-linearizing approximations in order to simplify calculation was then devised.[8] For prediction over an area approximately three-fourths of the size of North America, spanned in three dimensions by 1,083 grid points, it was necessary to use an external memory to augment the 1,024-word Williams' memory of the IAS machine. This was done first with punch cards and later by means of magnetic drum. A measure of the success of this model is that it forecast the occurrence of 90-mile-an-hour winds at the 3,000 foot level during the peak of development of the Thanksgiving Day storm.

Although we call our business numerical *weather* prediction, the only weather elements that I have spoken of are the wind and pressure fields. Actually, the predicted large-scale temperature field is also directly deducible. However, the predicted field of motion is only a necessary, though not a sufficient, prerequisite for predicting cloudiness and precipitation.

At least at the present, it appears that the large-scale field of precipitation can be predicted from a knowledge of the three dimension field of motion and the field of moisture.[9] Small-scale precipitation such as from individual thunderclouds is not predictable by the models described. In fact, our knowledge at this time of the

[4] J. G. Charney, "On a physical basis for numerical prediction of large-scale motions in the atmosphere," *Jour. Met.*, vol. 6, pp. 371–385; December, 1949.

[5] J. G. Charney and A. Eliassen, "A numerical method for predicting the perturbations of the middle latitude westerlies," *Tellus* (Stockholm), vol. 1, pp. 38–54; May, 1949.

[6] J. G. Charney, R. Fjörtoft, and J. von Neumann, "Numerical integration of the barotropic vorticity equation," *Tellus* (Stockholm), vol. 2, pp. 237–254; November, 1950.

[7] J. G. Charney and N. A. Phillips, "Numerical integration of the quasi-geostrophic equations for barotropic and simple baroclinic flows," *Jour. Met.*, vol. 10, pp. 71–99; April, 1953.

[8] Results to be published.

[9] J. C. Thompson and G. O. Collins, "A generalized study of precipitation forecasting, part 1: Computation of precipitation from the fields of moisture and wind," *Monthly Weather Review* (U. S. Weather Bureau, Washington), vol. 81, pp. 91–100; April, 1953.

mechanisms governing the formation, propagation, and dissipation of atmospheric disturbances of the scale of squall lines, tornadoes, and hurricanes, is so deficient that an adequate quantitative theory still remains to be formulated. The same may be said for the other end of the spectrum, namely weather changes over periods of the order of a month.

The Mathematical Problem

I would like to make some remarks regarding the nature of the mathematical problem involved and the workload that it presents to a high-speed computer.

In the models previously described, the prediction equations for an n-level model take the form of n-two (space) dimensional Helmholtz equations which must be solved simultaneously. For n equal to or greater than one, one of the equations reduces to Poisson's equation. Thus, for n equals one (the barotropic model) there is only a Poisson's equation to be solved. There are also n first-order linear equations in time which are solved by quadratures in a trivial calculation. The inhomogeneous terms of the Helmholtz equations involve Jacobian operations performed on the heights of a pressure surface and a quantity containing two-dimensional Laplacians of the height. Prediction for say 24 hours requires consecutive predictions for periods of half an hour, using the newly predicted field of motion as initial conditions for the next prediction. In general, the Jacobians must be calculated and Helmholtz equations solved for each short-period prediction. The Helmholtz equations, when transformed to finite difference form, are solved by means of a systematic over-relaxation scheme.

For the single level model, using a rectangular grid of 361 points, the machine performed 1,640,000 multiplications and divisions in order to yield a 24-hour prediction in one-hour steps. Approximately 13 iterations were required to reduce the relaxation residual to approximately one part in a thousand, so that the solution of the difference equation took 6 times as long as the formation of the Jacobians. This produced a 24-hour forecast in 48 minutes on the IAS computer. By rationalizing the code for maximum efficiency, and introducing a number of physical approximations, it was possible to reduce the time to 6 minutes. The two-layer model requires a little more than twice the time it takes for the barotropic prediction.

One can derive the equations for a general 3-dimensional baroclinic atmosphere, which in finite difference form with n internal vertical grid points reduces to the n-layered models when certain of the coefficients are fixed as constant. The general model requires the solution of a single three-dimensional Poisson equation in which the vertical second derivative has a variable and, in fact, nonlinear coefficient. It is quite likely that some version of such a model, with the further complexity of an earth's surface of variable height, will be used for subsequent experiments and ultimately for operational numerical weather prediction.

Requirements on Data

We tacitly assumed that somehow observational data in the correct form and appropriately processed are available to be used for initial conditions. The truth is that meterological data are not only far from suitable for numerical weather prediction but certainly are less than ideal for existing forecasting techniques.

Present methods for the measurement of meteorological elements, the transmission of the data, the processing at collection points, and finally the central recording have developed largely through patchwork and improvisation. Whenever a new need arose, its solution was sought on an individual basis and if some times this might have required redesign of the entire system, it was not economically feasible to do so. Under present practices, a piece of meteorological data passes through a "nondescript" gamut of processing—especially in the light of the requirements for operational numerical weather prediction.

In analyzing the meteorological data problem in general, one finds that in addition to the unnecessarily large time-lag between the taking of an observation and its ultimate usable form, there are many opportunities for errors to be introduced. In both instances the difficulty can be traced to the human element. I want to make it clear that the defects which I am pointing out cannot be corrected until adequate technological improvements are available—automatic instruments, better communications, high speed computers etc. The present system of weather data collection and processing may be the best that could be found until these communications and computer facilities came within reach.

The present framework governing the flow of data, because of its improvised mode of evolution, is in retrospect illogical, as illustrated by the following:

The analogue information from an instrument is evaluated and transformed into digital meteorological information and thereupon the data are entered into a station record manually; they are then put into automatic digital form by the manual punching of a teletype tape. When a collection of such data is received on another teletype tape, the digital message is converted to a written digital form by means of a teletype printer. *The tape is then discarded.* At perhaps 50 such central locations, some of the digital information on the digital sheet is then transcribed manually onto maps to be used for analysis and prediction. Others, who may wish the data for climatological purposes, will refer either to the original written station record, or to the teletype sheets, or to the plotted maps to again copy manually and perhaps rearrange the data. The paradox is that the data were in the most useful form when they were first placed on the teletype tape! Of course we must assume the tape was punched correctly at its source. Also, there is no guarantee that the message was not garbled in transmission, since there is no routine check. This example illustrates a number of weaknesses of the present system. Time has been wasted, manifold manual operations have been required, and errors have probably been introduced.

Collection

To give a concrete example of the time involved in collecting data, I have taken the following from the time schedule of the Weather Bureau-Air Force-Navy Analysis Center located here in Washington.

Upper air observations are taken simultaneously over the world twice a day at 0300 and 1500 Greenwich Meridian Time. The radiosonde apparatus takes about 20 minutes to get up to 400 mb of pressure or approximately 24,000 feet. While the balloon is still rising above this level, the lower portion of the sounding is evaluated and a teletype tape is punched within an hour. For the 0300 observation time, this means that a tape containing information from ground level to 400 mb (called the first transmission) is ready at 0400. For an area such as shown in Fig. 1, page 27, most such data are received centrally by 0630; however, it may be 2 or 3 hours later before all reports are in. The radiosonde apparatus reaches its normal top, which is 15 to 20 mb or approximately 90,000 feet, 80 or 85 minutes after the balloon is launched. Most of this information (the second transmission) is received by 0915 and again it may be two or three hours before it is all in. Thus for such an area it takes about nine hours for all of the data to be collected. If one wished all of the hemispheric data including Russia and Greenland, it might take another hour or two. This gets very close to the time for the next observation, and when one reflects on the fact that forecasts are highly perishable, we must conclude that the situation is highly incongruous.

There are approximately 160 radiosonde and 300 wind stations in North America, the North Atlantic, and the North Pacific. Of these, approximately half are in the continental United States. On the average, each radiosonde message (1st and 2nd transmissions taken together) contains 64 words and each wind message contains 52 words. A word consists of 5 decimal digits and a space. Thus for one aerological observation time, approximately 16,000 words need be collected. Transmitting linearly on a single teletype circuit at the present rate of 60 words per minute would consume $4\frac{1}{2}$ hours. In practice this is reduced by the use of multiple circuits. Some of the upper air winds are also transmitted at intermediate 6-hourly intervals. Thus roughly 50,000 words of aerological data for North America and the oceans are transmitted each day. It is of interest that this figure represents less than 10 per cent of the total meteorological wordage transmitted for the Northern Hemisphere. The bulk of this consists of observations of conditions at the surface of the earth taken each hour for use in airways forecasting.

Processing

Thus far we only have accounted for the time to collect raw information. These data are located at observation stations which are randomly distributed in the horizontal. The horizontal density of reports varies between large limits, very often falling below the necessary minimum, for instance in Canada and the oceans. However, predictions are still needed in these areas. The raw data contain small random instrumental errors and also physically real small scale variations which cannot be interpreted by the same mechanisms governing the large-scale variations. Such small-scale variations, instrumental or real, must be smoothed out. Furthermore, there may be gross errors in individual soundings which first must be detected and then must either be corrected or discarded. Finally, one would like an interpolated distribution of the quantities. The process which copes with these difficulties is called "analysis." The final form of the analysis is usually a representation of the variable in continuous graphical form mapped on a surface. An example is given in Fig. 2, page 28, where the observed heights of a pressure surface have been smoothed and interpolated. The contours are lines of constant height of a constant pressure surface. A good meteorological analyst takes years to develop a skill in making the necessary judgments to suitably process the data for prediction. The needs for numerical prediction have made the requirements for a good analysis much more critical and there is some question as to whether a good enough subjective analysis of the three-dimensional weather element distribution can be produced in a sufficiently short time in order that the analyzed data may be used for operational numerical prediction. Putting more men on the job is probably not the answer since all of the raw data must be scanned in a fully coherent manner and it is rather difficult to co-ordinate a group of men so closely. The natural question to ask is whether the judgments made by the subjective analyst are logical enough to be programmed. A psychoanalysis of the subjective analyst reveals that in general his decisions are rational rather than mysterious. He requires smoothness or continuity of the variables in time as well as in the three spatial dimensions, with weight depending heavily on the relative horizontal density of observations. His greatest skill lies in the smoothing and interpolation with respect to the horizontal distribution of observations, mainly as a result of years of conditioning in this type of analysis. Smoothing in time involves a crude short period forecast, and smoothing in the vertical in effect requires a re-evaluation of the original soundings with respect to horizontal and temporal smoothness. One finds that the subjective analyst is not always very thorough and consistent. As a matter of fact he sometimes even undoes the vertical consistency inherent in the original sounding since the raw data are essentially continuous in the vertical in contrast to their discreteness in the other dimensions. These deficiencies of the human analyst invariably can be traced to his inability to scan so much data as a function of four dimensions, even if the speed with which he can do this were not a factor. Now that I have made some disparaging remarks about the subjective analyst, and incidentally I am sure that many would argue violently, I would like to point out a skill which is not at all easy to duplicate by objective means —and that is the detection and possible correction of gross errors. The ability to do so is highly useful since every observation counts, and we would like to utilize an observation without its being misleading. For instance, the experienced subjective analyst can often make sense of garbled teletype messages.

For the past few months, we, at the Weather Bureau, have been performing experiments to determine the feasibility of objective analysis.[10] Preliminary results under controlled conditions indicate that in the case of horizontal smoothing and interpolation, the deviation of the objective analyses from the mean of 8 or 9 subjective analysis is no greater than the standard deviation of the subjective analyses. Considering that the subjective analyst has his greatest skill in a horizontal analysis, this is an extremely significant result. These experiments were performed over continental United States, where there was an adequate data density. Oceanic regions would present a greater challenge to objective analysis. Indeed, a preliminary experience indicates that a general objective analysis performed on a four-dimensional distribution of data with widely varying density will require as many logical as arithmetic operations. The problem is thus ideally suited for high-speed digital computers. It is hoped in the next few months actual tests will be performed toward development of a fully objective analysis system.

Visual Form

The most common visual representation of meteorological data is a geographical map upon which simultaneous observations of weather elements are plotted in coded form at the station locations. These are usually analyzed so as to give a field representation—for instance, Fig. 1. Both the plotting and, as has been pointed out, the analysis are time consuming and expensive, especially since this is duplicated at many forecast centers. For numerical prediction this need not be done if the raw data can be fed directly into a computer for objective analysis. Since the results of the objective analysis are interpolated-smoothed values of the quantities in a three-dimensional mesh, one can program a further interpolation to give the location of a contour of given value and then have a printer transcribe this information into a visual form. An example of the appearance of such a map is shown in Fig. 3, page 29. A similar representation can be given to results of numerical prediction. It thus appears that for large-scale meteorological work, one could completely dispense with the manual plotting and analysis functions for both the purposes of preparing initial data or presenting a visual picture of a prediction. It must be pointed out that for subjective small-scale prediction, the manual plotting and analysis routines or an appropriate substitute would still be necessary, but since this involves small local areas, the time consumed is not very large and duplication is nonexistent.

[10] Results to be published.

Dissemination

Up until 1946 almost all large-scale predictions formulated at an analysis center were distributed to district centers in coded forms via teletype. However, in the last 8 years facsimile has been utilized to transmit prognostic charts. The major deficiency of the present system is the time element. It takes 20 minutes for each chart to be transmitted. Often charts are detained for over an hour before there is an opening in the circuits. Possible solutions are additional circuits or higher speed facsimile.

There is an alternative which would require again a discrete communications system such as teletype. One could arrange for the results of an objective analysis or a prognosis in the form of isolines and value at mesh-points to be printed at the district centers. Signals could be sent directly from the computer or delayed by storage on a magnetic tape and then transmitted.

Some experimentation is already being carried out to speed the transmission of raw data for short-term airways forecasting. Under consideration is the selective broadcasting of blocks of data from a centrally located magnetic drum to given forecasting stations.

A Possible System

In the preceding analysis I have tried to outline present data handling practices and their deficiencies. In some instances, means for streamlining the process were self-evident. In other phases much experimentation must still be done. For instance it is essentially that an observational system, relatively inexpensively, give us adequate aerological data density in inaccessible regions, such as the oceans. It would be desirable that the analogue data from instruments be transmitted directly, in an attempt to eliminate all processing at observation stations. Higher speed communications systems must be utilized for the collection and dissemination of raw and processed information.

With a little imagination, one can lay out a possible system. We take license to incorporate devices which are yet to be developed. Referring to Fig. 4, page 30, we begin with automatic instruments which can remain unserviced for long intervals. Automatic surface devices have already been worked on, especially during the war when it was necessary to gain information behind enemy lines. Aerological soundings are a bit more difficult. One could consider floating vessels for which power is generated by wind, solar radiation or ocean shear currents. These vessels could also serve as microwave relay stations. Conceivably the radiosonde apparatus could be brought to a sufficiently high level by rocket and then dropped by balloon; or why not indirect sounding devices which do not depend on an instrument to pass through the atmosphere?

The analogue data from such instruments could then be relayed by the microwave network to collection points where the information is temporarily stored on a high-speed memory device such as magnetic tape.

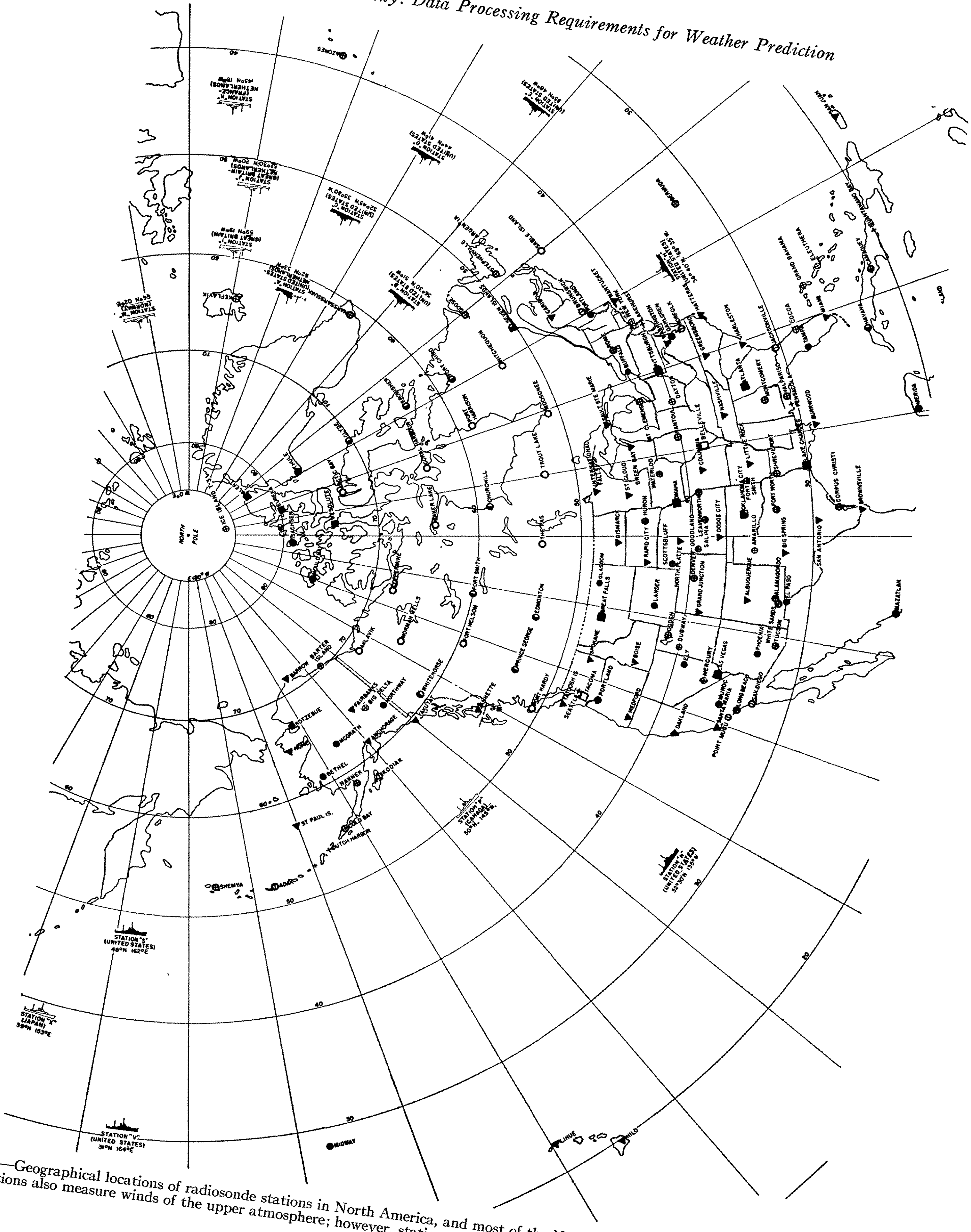


Fig. 1—Geographical locations of radiosonde stations in North America, and most of the North Atlantic and Pacific. The observation stations also measure winds of the upper atmosphere; however, stations which only observe winds are not shown. (Turn to read)

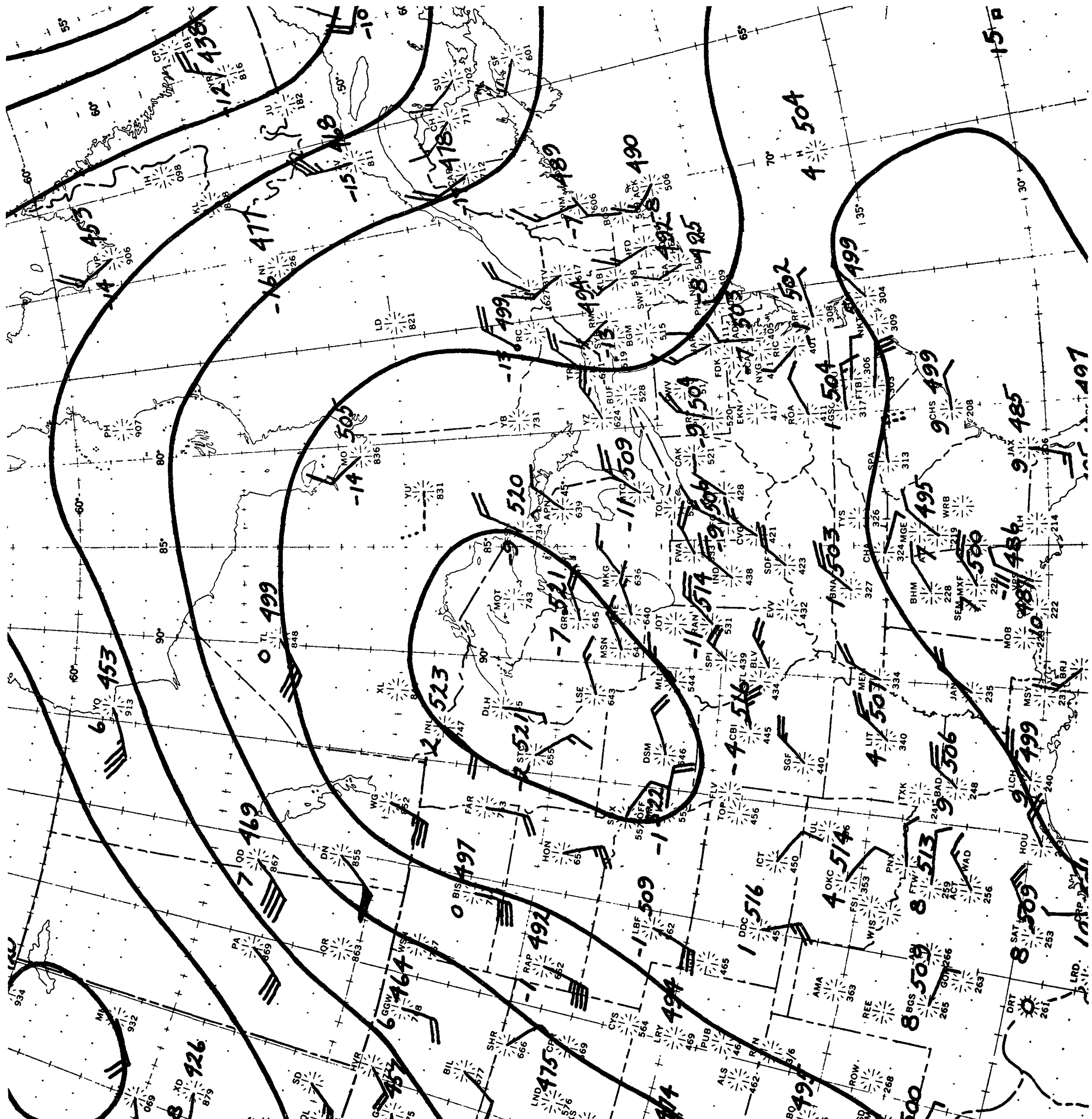


Fig. 2—Typical plot and conventional analysis of meteorological data. The plotted data refer to wind, temperature, and height of the 850 mb surface for 1500 Greenwich Meridian Time, November 5, 1953; the "analysis" consists of smoothed height contours of the 850 mb surface. (Turn to read)

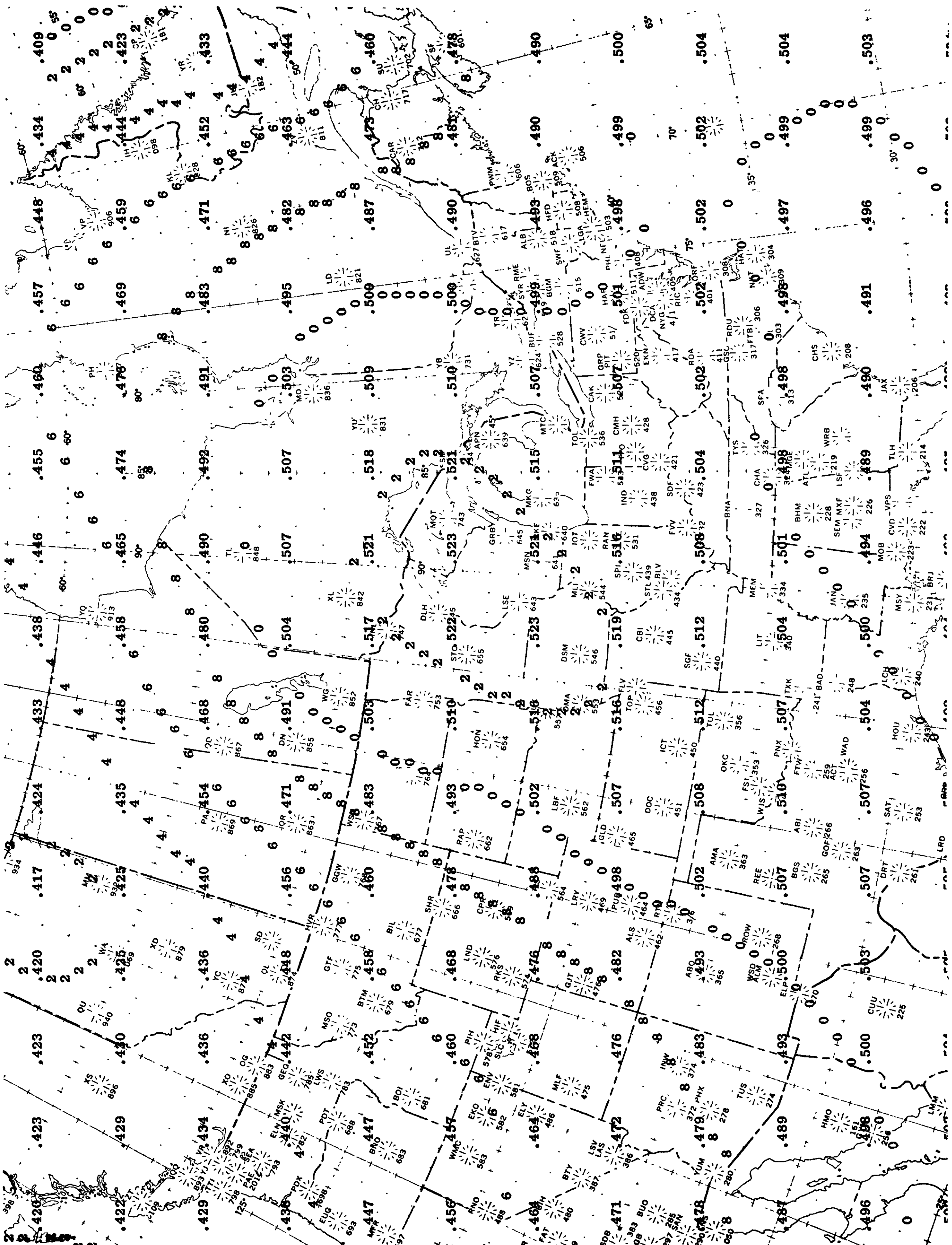


FIG. 3—Appearance of a computer controlled printing of an objective analysis of the data in FIG. 2. Shown are smoothed heights interpolated at grid points of a typical numerical prediction mesh, and contour lines. (Turn to read)

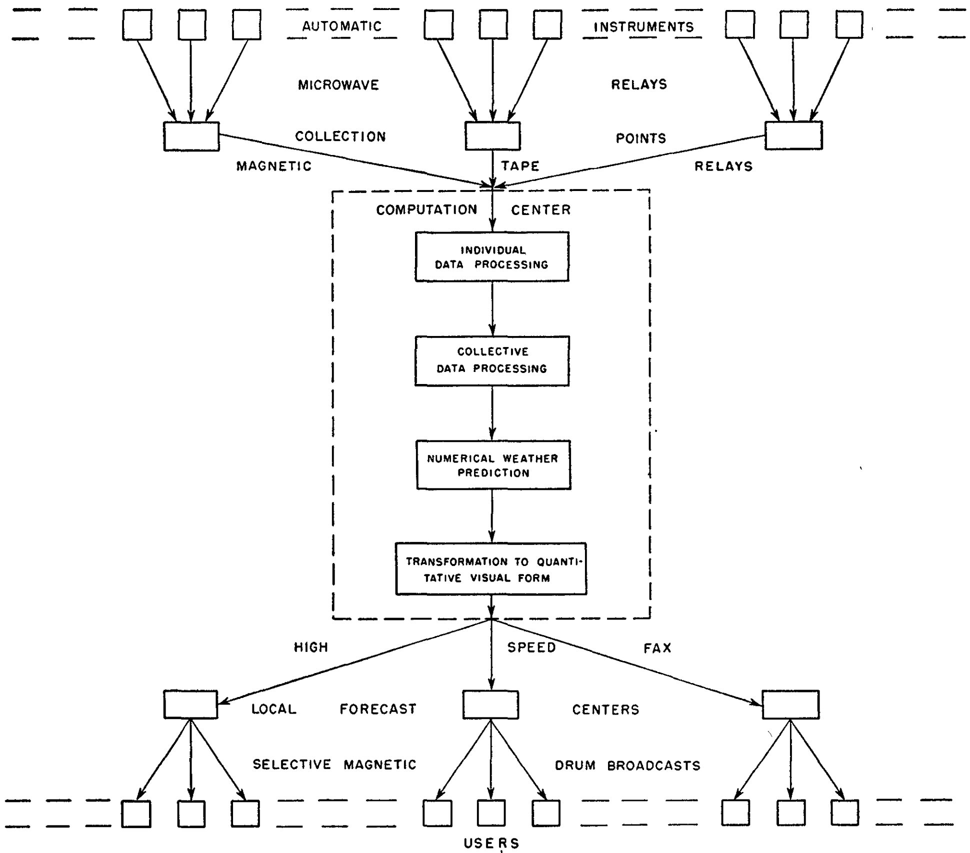


Fig. 4—A flow diagram of meteorological information in a possible system for data processing. The role of small-scale forecasting, extended-range forecasting, and climatological uses is not included.

From these points the information is funneled to a computation center. Here data are converted to digital form, checked for internal consistency, and processed as necessary on an individual basis. These data are then distributed to small-scale and long-range forecasting centers; for numerical prediction they are collectively processed by an objective analysis. A numerical prediction is then computed. The digital prediction is then transformed to visual form and transmitted by high-speed facsimile to local forecast centers.

Conclusion

It is quite apparent from the foregoing that our present system of procurement and processing of data is wholely inadequate. A critique of existing methods points out gross deficiencies and the fact that, in general, we are well behind the times. We have not taken advantage of the fruits of technological progress. All of this has been apparent to meteorologists for some time. In fact, there is presently under consideration a plan to study the data problem from the taking of an observation to its ultimate use. It is the possibility of the introduction of high-speed computing machinery into operational meteorology that has led us to the realization that the moment for reflection and change is urgently at hand.

Discussion

William P. Byrnes (Teletype Corporation): Would an increase in speed by a factor of ten satisfy your requirements for higher speed?

Mr. Smagorinsky: I assume that you mean communication speed. I would say a factor of ten would be highly desirable. It must be remembered that this probably would be reduced bv a factor of two almost immediately, if we impose the requirement that we must be able either to duplicate messages in order to be absolutely certain that we were receiving them as they were sent, or to incorporate some self-checking features in transmission. It is hard to say at this time whether ten would be fully adequate; but it certainly is in the right direction, a very large step in the right direction.

Mr. Byrnes: What type of transmitting facilities do you propose for higher speeds? Telephone channels? Microwave relays?

Mr. Smagorinsky: Frankly, I do not know; but I think this is a decision for communications engineers, to be made on the basis of requirements. Our requirements probably would not be much different from those of others. We would like the most reliable, fastest, and cheapest medium possible. This last item is very important as far as weather is concerned because, even at present, communications represent a very large amount of money in the running of the weather services.

Methods Used to Improve Reliability in Military Electronics Equipment

L. D. WHITELOCK†

RELIABLE, dependable, electronic equipment for industry and for the Armed Services has always been of prime importance for safety and for economy of operations. However, the rapid growth and the increasing complexity of electronic equipment have contributed heavily to a poor reliability record in military operations, with a resultant heavy drain on skilled personnel and funds for maintenance and supply functions. Indicative of this growth in size and complexity is the increase in electronic tubes used on a destroyer: from 60 in 1937, to 850 in 1944, and to 3,200 in 1952.

Top management in the Department of Defense became concerned about the urgent and complex problem of unreliable electronics equipment of the late 1940's and early 1950's. Reliability became a goal. Task forces of industry and armed services personnel were organized to prosecute a "Reliability Program" with the utmost application and vigor.

These task forces studied the causes of equipment failures in actual operation, and the causes of tube and component failures in these equipments. Broad factors affecting reliability were determined and the relative degree of unreliability contributed by each factor was evaluated. They found that each of the following factors affect electronic equipment reliability:

1. Specifications—Required: (a) to translate military characteristics into technical performance requirements; and (b) to standardize components.
2. Procurement—Covers both prime and subcontract procurement. This can have a great influence on quality.
3. System Design and Manufacture—System planning and design can have a profound effect on simplicity of the manufactured equipment.
4. Testing—Determines the ability of the equipment to meet the required performance.
5. Inspection—Verifies test results and tends to maintain quality of the product.
6. Packaging—Important for protection of equipment between time of shipment and time of use. Improvement in packing methods has made this a minor factor in the equipment reliability problem. However, defective packing can result in serious damage to equipment.
7. Transportation—A problem of time and transportation space. A reliability problem only if replacement equipment and maintenance parts are not available when needed. Very important in overseas shipments.
8. Supply—The ability to furnish equipment and maintenance parts, when and where needed, is important to reliable operation.
9. Installation—Proper installation by adequately trained crews improves reliability.
10. Operation—Adequate training reduces the misuse of equipment.
11. Maintenance—Good maintenance is essential to reliability.

Each service technical agency co-ordinates, manages, and controls technically the foregoing factors.

Task force studies on reliability were expanded to include the services of several commercial contractors. A mass of information was collected and analyzed. It was determined that vacuum tubes account for approximately 60 to 70 per cent of all equipment failures, which includes an indeterminate number of tube failures directly resulting from component failures. All other equipment failures were attributed to four major causes (vacuum tubes excluded) as follows:

1. Engineering (40 per cent)—Errors, omissions, negligence, and bad judgment of design engineers.
2. Components (30 per cent)—Components used in accordance with manufacturers specifications which were not abused, but which failed because of inherent defects.
3. Installation, Operation and Maintenance (20 per cent)—Equipments handled in a manner *not* in accord with prescribed procedures, without adequate procedures, or with incorrect procedures.
4. Manufacturing (10 per cent)—Equipment that is not built, tested, or inspected in accordance with proper specifications or in which poor workmanship exists.

The important lessons learned were that reliability was difficult to define in exact terms and that more effective teamwork between the military and industry was necessary to make any really significant improvement in reliability. This is because of the varied nature of the items causing failures. With such a major share of the nonreliability factor laid at the door of tube design and manufacture and to engineering of circuits and components, it was clear that these items must receive special attention. Therefore, specific programs were started to improve and strengthen the weakest points.

The first, and probably the most important, of these programs is aimed at the production of rugged, dependable vacuum tubes. In this program, the results of studies and statistical analysis of tube failures were used to select for improvement a number of tube types show-

† Bureau of Ships, Washington, D. C.

ing the heaviest use and the greatest failure rates. Extensive development work performed by several tube manufacturers has made a number of "premium" and "reliable" vacuum tubes available. These premium tubes cost more than their commercial counterparts; but they last longer when used in military service. As a consequence of their use, reductions in maintenance costs are obtained along with more reliable operation. More savings will be realized when current tube stocks are replaced with reliable tubes.

Qualified companies are developing new tubes in an advanced phase of the "reliable" tube program. These tubes will be more adaptable to automatic machine methods of tube assembly, to the use of new materials and to closer tolerances for all critical dimensions. Definite progress is being made along lines which will permit the quality of the tubes to approach the uniformity of machine tolerances. With present assembly-line techniques, the skill of the workers is the limiting factor.

The Reliable Component Program also receives close attention as an additional effort to improve reliability. Important objectives in this program are: (a) improved component specifications; (b) better inspection procedures; (c) specific approval of equipment component lists; and (d) the establishment of de-rating factors to be used by equipment designers for conservative design under certain conditions. Government laboratory tests are used as a basis for approval or disapproval of components on equipment lists.

The Reliable Sub-Assemblies Program for development of packaged and standardized sub-assemblies is being vigorously pursued for prospective use in relatively high-production items such as communication receivers.

Project Simplification is another part of the Reliability Program. Here operational and technical system requirements are reviewed in an effort to simplify the equipment. Specific steps have been taken to encourage laboratories and manufacturers to simplify circuits, equipment, and systems. Early examples under this program indicate that reductions of over 30 per cent in tubes and in components are often achieved.

A fifth and concurrent phase of the Reliability Program has been improvement in the General Specifications for Electronic Equipment, Naval Ship and Shore, with the recent issue of MIL-E-16400 (SHIPS) dated May 1, 1953. This specification is a revision of 16E4. The new specification includes definite goals for simplicity and reliability as follows: "The basic design objectives are that the equipment will meet the needs of Naval service and that the final product will reflect the utmost in simplicity, have maximum reliability consistent with the state of the art and be easy to install and maintain." The new specification includes many engineering guidance items resulting from the reliability studies. A summary of this specification is a valuable "check list" for project engineers, project supervisors.

Another method used to improve reliability is wide dissemination of technical information. The flow of reliability information is increasing from the Radio-Electronics-Television Manufacturers Association and from Department of Defense activities. Recent examples are: (a) "NAVSHIPS 91957—Reliability of Electronic Equipment," published by the Bureau of Ships; and (b) "Electronic Applications Reliability Review," published by the Radio-Electronics-Television Manufacturers Association. An excellent issue of the latter publication has just been released.

Field failure reports help improve reliability and reduce maintenance costs as much as any other measure used. Analyzing failure reports sometimes results in field change kits for existing equipment and design changes in new equipment. Both improve reliability. All failure data is now subjected to statistical analysis on high speed data processing equipment. The results are made available to all groups having a responsibility in connection with the design, manufacture or maintenance of the equipments involved. Further expansion of this effort is planned.

Reliability problems solved in the development of early large scale, high speed, digital computing equipments, have contributed substantially to improved engineering techniques. Many of the engineering guidance items from General Specification MIL-E-16400(SHIPS) are very applicable to computers and some are an outgrowth of computer work. Based on experience, I suggest close attention to the following in computer design:

1. The use of JAN or MIL or equivalent tubes, resistors, condensers, and other components, with the use of proper de-rating factors where listed in MIL specs.
2. Conservative mechanical and electrical design.
3. Good workmanship.
4. Air conditioning, including both heating and cooling in some cases, particularly where transistors are involved.
5. Marginal testing.
6. Built-in test center.
7. Unit construction.
8. Simplified and standardized circuits.
9. Spare chassis and special test fixtures.
10. Maintenance accessibility.
11. Adequate test and inspection.
12. Trained installation crews where the size of the equipment warrants.
13. Adequate maintenance training.
14. Adequate supply of maintenance parts, material.
15. Adequate operational training.

In summary, the methods used to improve reliability in military electronic equipment are:

1. Effective analysis of the factors involved.
2. Apply corrective measures necessary for the improvement of each factor.
3. Continuous dissemination of pertinent technical information.

4. Expanded research and development in the areas found technically deficient.

These methods are, of course, equally applicable in industry.

In conclusion, I believe that electronics equipment reliability can, and will, be tremendously improved by the increased effort now focused on this problem. I believe that the degree of acceptance of computers and other new products by industry and the military will depend largely on the degree to which reliability has been built in. Better tubes, better components, and progressive engineering are the key building blocks.

Discussion

C. T. Schaedel (Convair, Fort Worth): Please restate the MIL spec number and another reliability report which has been mentioned.

Mr. Whitelock: The general specification that I mentioned was General Specifications for Electronic Equipment, Naval Ship and Shore, MIL-E-16400 (ships). It is a military spec used by the Bureau of Ships. If any of you desire copies, I believe that a letter on your business letterhead addressed to the Chief of the Bureau of Ships, Washington 25, D. C. would receive favorable attention. The second publication that I mentioned was the Electronics Application Reliability Review, which has published two issues to date. It can be obtained through the engineering office of Radio Electronics Television Manufacturers Association at 500 Fifth Avenue, New York 36, New York.

J. R. Harris (Bell Telephone Laboratories): What is your opinion of the value of built-in error-detecting equipment and of built-in error correcting equipment?

Mr. Whitelock: Both can be valuable, depending on where and how they are used. In computers, error detecting equipment is particularly useful in those cases where you plan to use the equipment for logistics or accounting purposes. There are some types of applications where running the problem twice is perhaps a better solution. Error correcting equipment is used in a number of ways, but not necessarily in computers. Without getting into a general discussion. I can say that it is a very useful tool, for instance, in communication receivers to automatically control oscillator frequency.

George J. Christy (State of California, Department of Education): I want to know whether your surveys covered all branches of the government that are using electronics equipment and where can material be obtained concerning your findings?

Mr. Whitelock: The material that I reported on here was essentially Department of Defense, and was a summary, very briefly given, of work that started out as a joint action of industry and the Research and Development Board. Not all of it is available for relase. In general, the Electronics Application Reliability Review that I mentioned earlier will probably give you most of the information you want, if you can get Issue 1 and Issue 2, and the supplements to Issue 2.

David Thorndike (Financial Publishing Company): What per cent down time does tube and component failure account for? What is the per cent of operating time and of total available time?

Mr. Whitelock: I do not know quite how to answer this question. Applying it to computers, which I presume Mr. Thorndike means, it is very difficult to put tubes and components together and get a real percentage. However, I would estimate that 95 per cent of failures are attributed to this combination. Most of the large scale computers are now operating at about 85 per cent, or a little better, good operating time, which means that the equipment is available for use 85 per cent of the time. Roughly ten per cent of the remaining time is used for various marginal checks. Components and tubes account for probably five per cent unscheduled down time. In other military electronic equipment where marginal checks are not used as widely yet as they are in some of the computers, the unscheduled down time due to tube and component failures may be quite high.

Digital Computers for Linear, Real-Time Control Systems*

RALPH B. CONN†

Summary—**A digital computer operating in a real-time control system is an information-processing device. Heretofore, special-purpose, general-purpose computers have been designed for use in these systems. This paper begins by considering the basic equation for discrete, linear, real-time control systems. The operations that must be performed are enumerated and described. Several methods for accomplishing each operation are discussed, pointing out the adadvantages and disadvantages of each method. In conclusion the weighting-function computer is compared with the special-purpose, general-purpose computer and the digital-differential analyzer.**

* This paper presents results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. DA-04-495-Ord 18, sponsored by the Department of the Army, Ordnance Corps.

† Formerly with the Jet Propulsion Lab., Calif. Inst. of Tech., Pasadena, Calif., now with The Ramo-Wooldridge Corp., Los Angeles, Calif.

Introduction

THIS PAPER DESCRIBES the logical organization of digital computers for use in control systems. Most of the digital computers used or contemplated for use in real-time control systems have been operated basically in a general-purpose manner.[1] After each operation is completed, the control obtains a new order from the memory, decodes it, and causes the arithmetic element to perform the proper operation on the correct number. The general-pupose computers used in control systems are distinguished by their

[1] E. C. Nelson, "A digital computer for airborne control systems," Trans. I.R.E. PGEC, vol. 1, pp. 2–5; December, 1952.

shorter word length and lack of ability to modify their orders.

The logical organization of digital computers allows great flexibility in the operation and adds greatly to the utility of the general-purpose computer. However, it will be observed that this flexibility is not needed in many control applications where in most cases a very restricted class of equations is considered. In such instances, tailoring the computer's design to the equations is more effective than making the equations fit the computer, a process which does not use the full potential of the general-purpose computer and, moreover, asks the computer to perform tasks for which it was not designed. As a result, when general-purpose computers are used in control systems, high computing speeds become necessary, placing stringent requirements on the electronic circuitry.

In the following discussion the philosophy of a computer designed to solve control equations and the components of several possible configurations will be explored.[2] These computers can be used in both open-loop and closed-loop control systems.

Equations To Be Solved

The computers to be described solve one type of equation. Fortunately, this is a very general type of equation. In fact, it is the most general type for linear, time-invariant digital control systems.[3] The form of the equation is

$$o_k = \sum_{j=0}^{M} a_j u_{k-j} + \sum_{j=0}^{N} b_j v_{k-j} + \cdots + \sum_{j=0}^{P} c_j w_{k-j} \tag{1}$$

where u_k, v_k, $\cdots$, w_k are the several input variables; a_j, b_j, $\cdots$, c_j are the appropriate weighting constants; M, N, $\cdots$, P are finite integers; and o_k is the present value of the output. If there is only one input variable, e.g., u_k, (1) reduces to

$$o_k = \sum_{j=0}^{M} a_j u_{k-j}. \tag{2}$$

Equation (2) is the discrete equivalent of a filter-weighting function operating on the u_k. Here the digital computer acts as a simple filter. If there are several input variables, (1) is appropriate. This type of equation arises in open-loop control systems. An equation taking the form

$$o_k = \sum_{j=0}^{M} a_j u_{k-j} + \sum_{j=1}^{N} b_j o_{k-j} \tag{3}$$

describes a closed-loop system.

[2] J. M. Salzer, "Treatment of Digital Control Systems and Numerical Processes in the Frequency Domain," Doctor of Science Thesis, Massachusetts Institute of Technology, Cambridge, Mass., 1951. Salzer devotes a paragraph to a description of a computer of the type to be discussed in this paper. While many of the features to be described in this paper were not mentioned by Salzer, he should receive credit for recognizing the possibilities of such a computer.

[3] Salzer, *loc. cit.*, p. 68.

Thus, the object of this paper is to demonstrate how (1) can be mechanized so that a minimum of hardware operating at a low computing speed is attained.

Necessary Operations

Since the term in (2) is basic to (1) and (3), only (2) is considered initially. The solution of (2) requires that the following items be provided: (1) a representation of the u_{k-j}, (2) a representation of the a_j, (3) a method for multiplying u_{k-j} by a_j simply and rapidly, (4) a means for summing the $a_j u_{k-j}$ products, and (5) an easily mechanized scheme for sequencing the u_{k-j}. Of these, only items (3) and (5) provide much difficulty.

The method used for multiplying is very important. In ordinary serial machines, multiplication takes much longer than do other arithmetic operations. It would be desirable to accomplish multiplication in the same length of time as that taken by other operations. This problem is discussed in the following section.

Sequencing the u_{k-j} is equally as important as multiplication. This operation is necessary because, with each new input, past data become older with respect to the present computation and must receive different weight. The sequencing is complicated by the fact that real-time control is desired and that the sequencing should be easily mechanized. In the ordinary serial machine, sequencing is accomplished by reading the number into the arithmetic element and, sometime later, transferring it to the appropriate place in the memory, a process which can result in a large amount of wasted time. The solution to this problem will be discussed later.

Multiplication by Weighting Constants

The first multiplication method uses a binary representation of the u_{k-j} with a precision of one part in 2^n-1. The a_j are represented as binary numbers but are restricted to the values $\pm 1/2^r$ where $r=0, 1, 2, \cdots, n+1$, permitting the multiplication process to be greatly simplified. Using these restricted values of a_j, multiplication is accomplished by shifting the u_{k-j} r places to the right. Since the smallest admissible a_j is $1/2^{n+1}$, it is never necessary to shift u_{k-j} more than $n+1$ places to the right, making possible multiplication in one word time. This type of multiplication admits only the coarsest of weighting constants and probably is not adequate for most control problems.

If computing time is available, the process of multiplication by shifting can be extended by multiple processing of the inputs. Thus, in order to obtain a weighting constant of 13/16, the appropriate input might be multiplied first by 1/2, next by 1/4, then by 1/16, and finally the three products summed. In general, if p is the number of processings allowed, it is possible for the a_j to be within $\pm 1/2^{2p-1}$ of any value $-1 \leqq a \leqq 1$. It is obviously possible to carry this process to the point where a complete multiplication is performed.

Another method of accomplishing multiplication in a short time is to perform the operation in parallel. This is a very expensive solution, requiring n times the circuitry used to multiply an n digit number serially. However, if a time limitation excludes multiple processing or if the incremental scheme to be described is undesirable, parallel processing may be the only solution.

If the inputs can be scaled so that the only possible input changes between computations are ± 1 unit or zero, multiplication becomes addition:

$$o_k = \sum_{j=0}^{M} a_j u_{k-j}, \qquad o_{k-1} = \sum_{j=0}^{M} a_j u_{k-1-j}, \tag{4}$$

$$\Delta o_k = o_k - o_{k-1} = \sum_{j=0}^{M} a_j (u_{k-j} - u_{k-1-j})$$

$$= \sum_{j=0}^{M} a_j \Delta u_{k-j}. \tag{5}$$

Thus Δo_k, the change in the output, is merely the proper combination of the a_j. This type of data processing offers the interesting possibility of either synchronous or asynchronous operation; i.e., a computation is made only when the input changes by ± 1 unit. Asynchronous operation offers the possibility of a variable-length weighting function. This incremental input system has the usual disadvantage of such systems: it is possible for the input device to lose the absolute level and thus produce erroneous input data. If a sufficiently reliable input device is available, incremental multiplication offers the maximum answers per second with the least arithmetic circuitry.

Sequencing the Variables

In this discussion it is assumed that a serial-storage unit is available. Since magnetic drums are widely used, the sequencing ideas are explained using a magnetic drum-storage unit as an example.

The rows of Fig. 1 represent the contents, at successive word times, of a portion of one track on a magnetic drum. R_1 and R_2 are read heads spaced two-word lengths apart, and W is a write head. For the time being, the portion of the drum circumference that this information occupies is unspecified. For definiteness of explanation, assume that the u_k are to be multiplied by the a_j using the restricted-constant multiplication method.

To start a computation, u_{k-m} is read into a shifting register through read head R_2. Using write head W, u_{k-m} is also rewritten on the drum. At the end of this readout, the information on the drum track has the physical position shown in the second row of Fig. 1; i.e., a_m is in position to be read from read head R_1. If a_m has been recorded as a series of pulses and is read from R_1 and applied to the shift bus of the register containing u_{k-m}, at the conclusion of the a_m readout the auxiliary shift register contains $a_m u_{k-m}$. The third row of Fig. 1 shows that it is now possible to read u_{k-m+1} from R_2 into the shift register and also to rewrite it on the drum using write head W. At the same time, $a_m u_{k-m}$ is read out of the shift register through the adder and added to any previous products or constants stored in the au product register, and the new sum is placed in the au product register. The reading, rewriting, and processing of the u_{k-j} continues until u_{k-1} has been processed. At this point u_k is read from the input device into the shift register and also applied to write head W. The weighting constant a_0 is then read from R_1, forming $a_0 u_k$, which is added to $\sum_{j=1}^{m} a_j u_{k-j}$ in the au product register.[4]

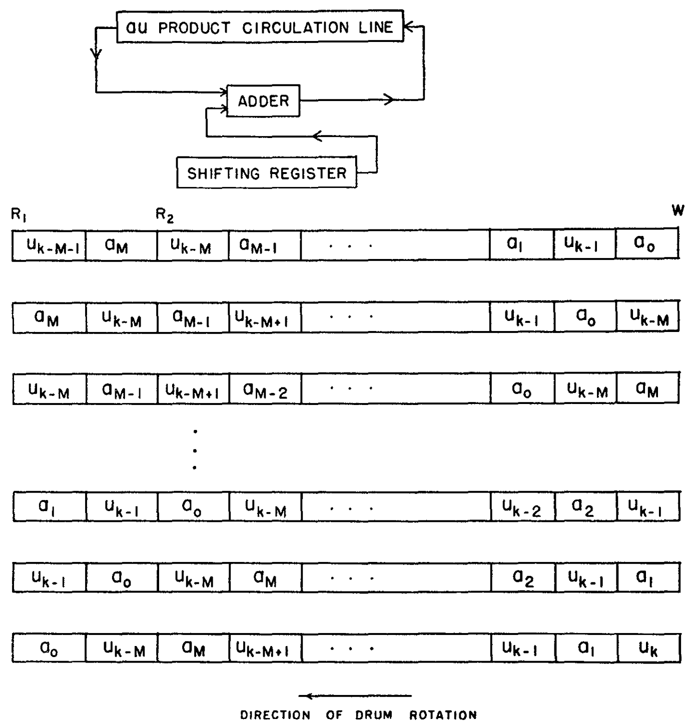


Fig. 1—Sequencing of the u_{k-j}.

The data stored on the drum track have now received a coarse weighting and have been sequenced in preparation for the k plus first computation. If further weighting of the u_{k-j} is necessary, it can be accomplished by using a_j stored on other parts of the magnetic drum. Although the sequencing method was explained using the restricted constant type of multiplication, the same sequencing method is applicable without regard to the multiplication method.

The $a-u$ storage line, as already described, can occupy any fraction of the drum's circumference. It is convenient, however, to have the $a-u$ storage line occupy a unit fraction of the drum's circumference. Thus, the drum speed may be used as the reference timing, a step which eliminates rewriting of a_j as the data processing proceeds, thus minimizing rewriting errors.

[4] Negative values of a_i are handled by storing the sign of a_j in the last pulse position of a_{j+1}. This sign pulse causes the complement of u_{k-j} to be read into the shifting register. Shifting the complement of u_{k-j} leaves $-|a_j|u_{k-j}$ in the shifting register.

So far, only the weighting of past and present inputs has been considered. If weighting of past outputs is necessary, as shown in (3), the sequencing is accomplished as before. However, the spot occupied by o_k will be vacant until the computation is completed. At the time o_k is being read out of the machine, it is read onto the appropriate spot of the drum.

Multiplication of Inputs

If the control system under consideration requires that an input be formed by multiplication of two variables, two short cuts are available. One method is to take the logarithm of each variable, add these two, and form the antilogarithm. The formation of the logarithm and the antilogarithm offers a saving in time over conventional multiplication but increases circuitry complexity. Obviously neither variable can be allowed to go to zero. Another method exists only if incremental inputs are available. This type of multiplication results from a difference approximation to the equation:

$$\int \frac{d(xy)}{dt} dt = xy = \int y \frac{dx}{dt} dt + \int x \frac{dy}{dt} dt. \quad (6)$$

The approximation is

$$(xy)_k = \sum_{j=0}^{k} x_j \Delta y_j + \sum_{j=0}^{k} y_j \Delta x_j. \quad (7)$$

If only changes in xy are desired as inputs to the computer, (7) can be simplified to

$$\Delta(xy)_k = x_k \Delta y_k + y_k \Delta x_k. \quad (8)$$

This method also offers circuit complexities, but a price must be paid for formed inputs.

Comparison of Weighting-Function Computer with Special-Purpose General-Purpose Computers

A special-purpose, general-purpose computer, as described earlier, was programmed to execute a typical control problem involving present and past inputs and past outputs. The weighting-function computer used for comparison was one that used multiple processing of data to accomplish complete multiplication. The results are summarized in Table I.

Table I does not make apparent the fact that the weighting-function computer does not have to be programmed as does the special-purpose, general-purpose machine. Thus if extra space is available in the $a-u$ storage lines, a weighting function can be lengthened or shortened by giving the a_j the appropriate values.

Comparison with a Digital-Differential Analyzer

A detailed comparison of a weighting-function computer and a digital-differential analyzer has not been made. This comparison would require designing a digital-differential analyzer specifically to solve equations of the type shown in (3). However, a few general remarks will be made.

Fundamentally both computers solve difference forms of differential equations. However, the differential analyzer must be programmed. This requirement provides little difficulty if the control equations arise from a continuous analysis of the system; but, if the input data are considered to be discrete, as they actually are, programming can become a chore.

Another disadvantage of the differential analyzer is that it is limited to incremental inputs with their attendant difficulties. Moreover, the differential analyzer must have its method of integration built in, whereas the weighting-function computer allows a selection of integration methods, if integration is involved.[5]

The digital-differential analyzer is much better suited to control problems requiring co-ordinate transformations than are the weighting-function computers. The differential analyzer can also handle nonlinearities, whereas in its present form the weighting-function computer cannot.

TABLE I

Comparison of Weighted-Function Computer with Special-Purpose, General-Purpose Computers

	Special-Purpose, General-Purpose Computer	Weighting-Function Computer
Total tube count (excluding power supply and input-output equipment)	170	90
Number of heads needed for magnetic drum	5	9
Required drum speed (rps)	280	60
Computing speed (kc/sec)	341	36.8
Computing efficiency (%)[1]	60	100

[1] Computing efficiency is defined as the ratio of time during which actual computing is occurring to the total time between computations.

Reliability Requirements

Digital computers used in real-time control systems *sometimes* demand reliability requirements significantly different from those of general-purpose machines. For instance, control problems arise where the computer must operate error free for a period as short as five minutes. This type of situation makes possible more frequent checks than with the large-scale, general-purpose applications. However, there are other types of control problems where several hours of error-free operation are required, and the penalty of an error might run into hundreds of thousands of dollars.

The computer discussed in this paper should offer a significant increase in reliability because of the reduction of the number of tubes and the lowering of both drum-rotation speed and pulse-repetition rate. As yet, no actual operating experience is available.

[5] M. Palevsky, "The design of the Bendix digital differential analyzer," Proc. I.R.E., vol. 41, pp. 1352–1356; October, 1953.

Conclusion

The basic elements of several possible computers for use in linear, time-invariant, real-time control systems have been described. These computers offer significant savings in the number of vacuum tubes required and allow the computer pulse-repetition frequency to be lowered appreciably. These results point to the fact that, rather than adapting the general-purpose computer to specialized tasks, the engineer should design the computer for the specific job it is to perform.

Acknowledgment

The author wishes to acknowledge the benefits derived from discussions with J. J. Burke, G. J. Gleghorn, W. A. Koehn, and N. L. Kreuder during the course of this investigation by the Jet Propulsion Laboratory.

The MIT Magnetic-Core Memory*

WILLIAM N. PAPIAN†

Introduction

ONE RECENT DEVELOPMENT which is significantly raising the reliability of today's high-speed automatic digital computer is the multicoordinate magnetic-core memory. Two banks of 32 by 32 by 17 magnetic-core memory have been in full-time operation in the Whirlwind I Computer for some months. A description of the units and of the tests and operational data available on them will be preceded by a short review of operating principles of this type of memory.

Operating Principles[1]

Each binary digit is stored in the magnetic field of a small, ring-shaped, ferromagnetic core. Two aspects of the core's rectangular flux-current characteristic are utilized:

a. The flux remanence of the core is utilized for the storage operation;[2]

b. The extreme nonlinearity of the flux-current characteristic is utilized to advantage in the selection operation.[3,4]

Fig. 1 shows the flux-current loop for a ferrite core. The remanent flux points are arbitrarily designated as ZERO and ONE. Note that the loop is sufficiently nonlinear so that the application of $I_m/2$ cannot switch the core, whereas the full I_m can. Fig. 2 illustrates how this nonlinearity may be used to select one core out of many by the coincidence of two half-currents in a 2-co-ordinate scheme. The extension to three co-ordinates may be accomplished by stacking planes like those of Fig. 2 behind each other and connecting respective x and y co-ordinate lines in common to obtain a "volume" of cores as sketched in Fig. 3, following page.

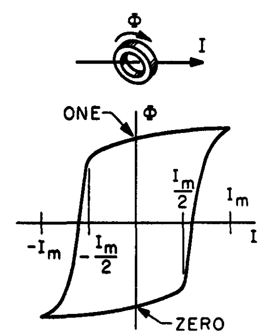


Fig. 1—Flux-current characteristic of ferrite toroid.

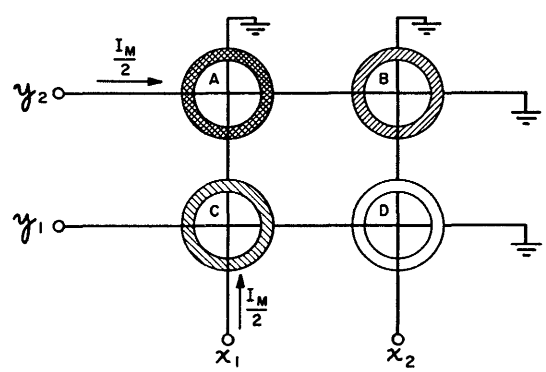


Fig. 2—Two-co ordinate array.

The application of a half current to the co-ordinate x_l results in the half excitation of a "selection plane" through the volume. The same is true for the co-ordinate y_m, and the result is full-current excitation of the line of cores at the intersection of these two selection planes. The internal memory for a parallel type of machine might well resemble Fig. 3, and the selected line of cores might well represent the selected memory register, or word. A readout or sensing winding threaded through

* The research in this document was supported jointly by the Army, Navy and Air Force under contract with the Massachusetts Institute of Technology.

† Staff Member, Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

[1] W. N. Papian, "A coincident-current magnetic memory cell for the storage of digital information," Proc. I.R.E., vol. 40, pp. 475–478; April, 1952. (Also "A Coincident-Current Magnetic Memory Unit," Master's thesis, E.E. Dept., Massachusetts Institute of Technology; August, 1950.)

[2] Harvard University Computation Laboratory, "Investigations for design of digital calculating machinery," *Progress Reports*, 2–6 (particularly No. 2); August, 1948–November, 1949.

[3] J. W. Forrester, "Digital information storage in three dimensions using magnetic cores," *Jour. Appl. Phys.*, vol. 22, pp. 44–48; January, 1951.

[4] Jan Rajchman, "Static magnetic matrix memory and switching circuits," *RCA Rev.*, vol. XIII, pp. 183–201; June, 1952.

every core in each *xy* plane, or digit plane, would bring out the signal representing the stored digit. This part of the read operation is destructive, and the word must be rewritten.

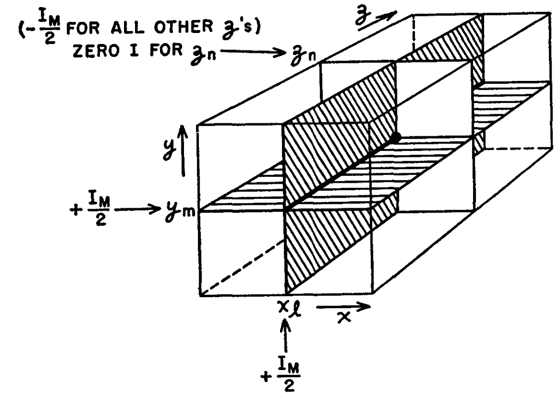


Fig. 3—Three-co-ordinate selection.

For the rewrite part of the cycle the selection technique remains the same, except that the half currents on the selection planes are now in the write polarity, which would result in the writing of ONE's into all the cores of the selected register; this writing is controllable for each *xy*, or digit, plane by the use of a digit-plane winding on which may be applied a half-current of an effective polarity opposite to the write currents. The presence of this "inhibit" current in any digit during the write operation leaves a ZERO; absence of the inhibit current leaves a ONE.

R. R. Everett of Massachusetts Institute of Technology has shown that these techniques may be extended into any number of co-ordinates but that the 2-co-ordinate read and 3-co-ordinate write system just described is one of the most desirable for the Whirlwind type of machine.

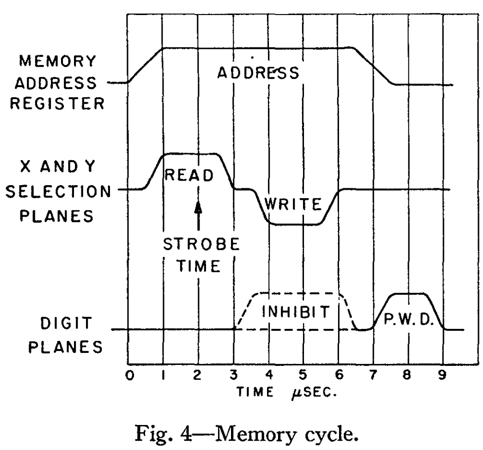


Fig. 4—Memory cycle.

Description of Memory

The capacity of each memory bank is 1024 registers, with 16 digits (plus 1 parity digit) per register. The basic operating mode, or cycle, consists of setting the memory-address register to the new address and applying the read-current pulses, followed by the write currents for rewriting the information just removed. The information is stored in a memory-buffer register. The speed of the machine may be judged from the timing diagram (Fig. 4). Note that the read-rewrite time, or cycle time, is approximately 9 μsecs and that there are no restraints on how frequently this cycle may be applied to the memory. It is capable, therefore, of a basic repetition rate of over 100 kc. Note also that the information can be available to the machine approximately 2 μsecs from the beginning of the cycle.

Block Schematic

Fig. 5 shows a block schematic of one bank of memory. Each half of the binary address in the Address Register is translated to a 1-out-of-32 selection by a crystal-diode Matrix and sets up a pair of "AND" gates for *x* and a pair for *y*. The Read Flip-Flop forms a 1.5-μsec pulse and sends it to the two selected Read Drivers which supply the 0.45-ampere currents to two selection planes. The output signal voltages from each digit plane are amplified in the Sense Amplifiers and applied to "AND" gates which are strobed at the optimum time by a short (0.1 μsec) pulse. Pulses representing ONE's then go off to set the Buffer Register to the just-extracted number. At the end of the read currents the rewrite part of the cycle starts in the same manner, except that the write currents have to be safely overlapped by the inhibit currents at those digit planes where ZERO's are to be written. This is accomplished by having the "on" time of the Inhibit Flip-Flop overlap slightly that of the Write Flip-Flop. Short (1 μsec) currents may be applied to all digit planes after the rewrite; they are called Post-Write Disturb (PWD) currents and are used to improve the ONE-to-ZERO signal ratios under certain conditions. The PWD Flip-Flop forms this pulse and applies it to all 17 of the Digit-Plane Drivers through "OR" inputs.

The Cores

The cores are made of General Ceramics material MF-1326B. The first bank contains their core size F-291 which has an outside diameter of 90 mils. A smaller core was used in the second bank; this is F-394, 80 mils in outside diameter. Single-turn switching currents are approximately 950 and 850 ma respectively, and single-turn output voltages (at optimum strobe time) are about 0.1 volt. Switching time, under these conditions, is approximately 1.2 μsecs.

One of the largest problems in the building of a memory of this type is the procurement of large numbers of uniform cores. Core selection was made on the basis of a series of pulse tests, approximately four per core, and resulted in a yield for the first bank of approximately 30 per cent of those shipped to us by the producer. (Yields have been improving materially since this first run.) The selection criterion was fundamentally that of an upper and lower limit on the voltage output from each core when the core was excited by a sequence of current pulses devised to resemble computer operation.

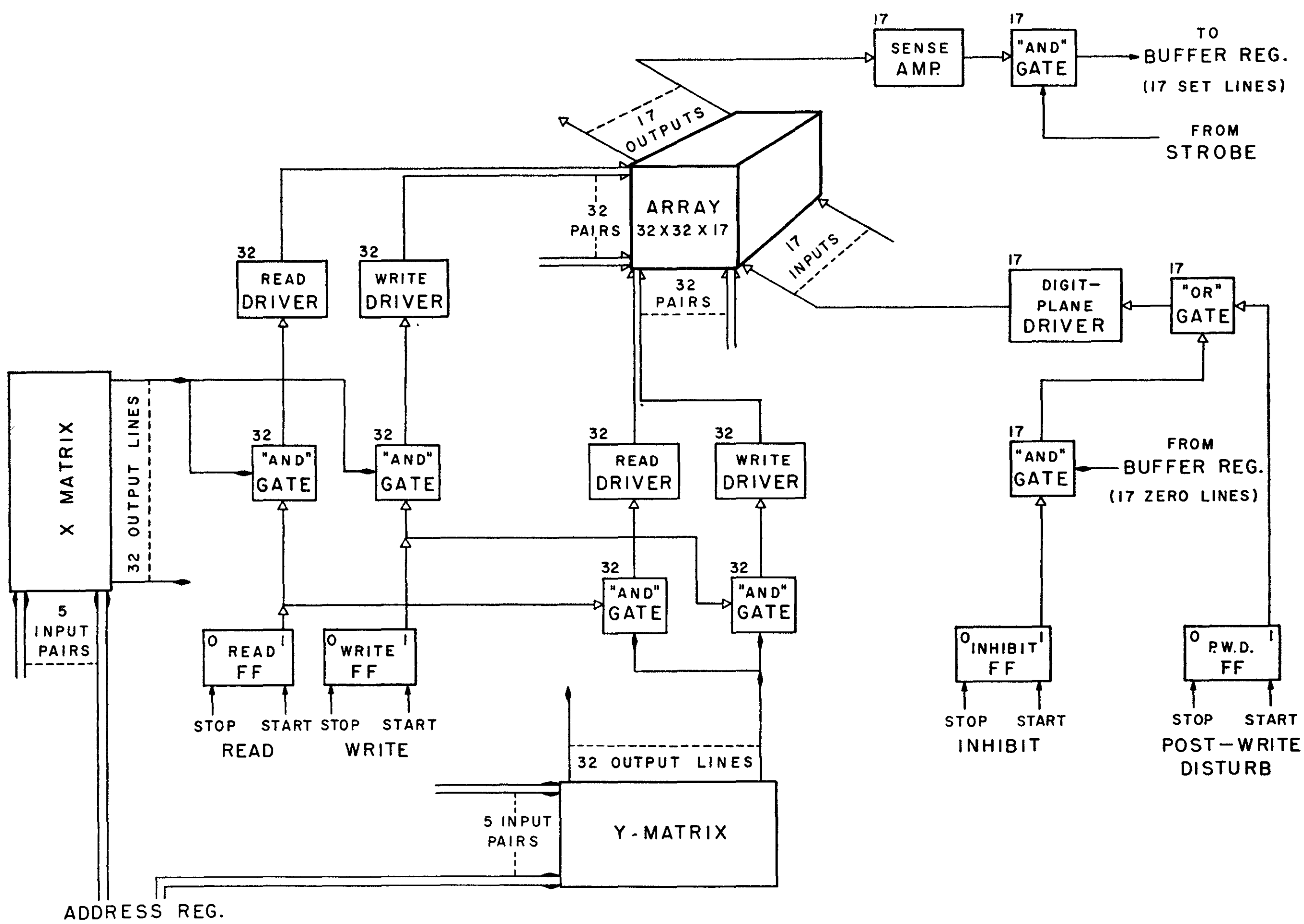


Fig. 5—Memory block schematic.

Fig. 6 shows typical output-voltage pulse shapes, the nominal limits within which cores were considered acceptable, and the strobe time at which these amplitudes were taken. The horizontal limit lines are at 90 and 120 mv, total pulse length is about 1.2 μsecs, and the vertical line showing the strobe time is about 0.5 μsec from the start of the pulse.

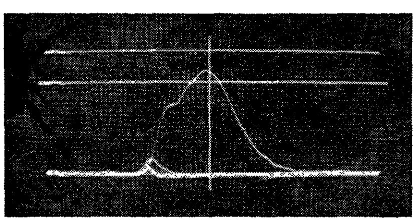


Fig. 6—Test-core outputs.

Basic-Circuit Types

The read and write currents for the selection planes of the memory are supplied directly from vacuum-tube plates. A single type-6080 vacuum tube, with its sections paralleled, is used to drive a given selection plane in the read direction. Another such tube drives the same plane in the write direction. The control grids of the 6080's are driven through 6BL7 amplifiers from the crystal-matrix output lines. The cathodes of all of the 6080 tubes in the x-read group are connected together, then through a large resistor to a negative-voltage supply. The cathodes of the three other groups of 6080's (x-write, y-read, y-write) are all connected in a similar manner. Each group of cathodes is normally held at a relatively high potential by a power amplifier and is allowed to drop at the proper time. Thus, each 6080 acts not only as a cathode follower but as the logical "AND" gate shown separately in the block schematic. The large amount of degeneration caused by the high common-cathode resistor compensates for nonuniformity and aging changes in the characteristics of the tubes. As a result, selection-plane currents remain within very close limits (plus or minus 4 or 5 per cent).

The digit-plane driver consists of a 6080 dual triode driven from two amplifier stages and incorporating sufficient negative feedback from the output to the input to keep the current amplitude within plus or minus 5 per cent over expected tube, component, and power-supply variations.

The output signal from the sensing, or read-out, winding is linearly amplified from the 100-mv level up to approximately a 30-volt level in a single-sided, ac coupled, wide-band feedback amplifier. The signal is then rectified and applied to the suppressor grid of a 7AK7 gate

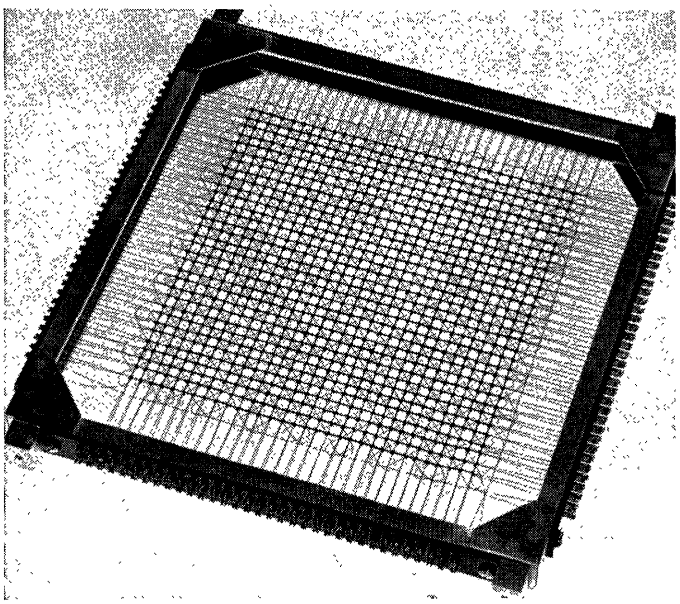


Fig. 7—32-by-32 plane.

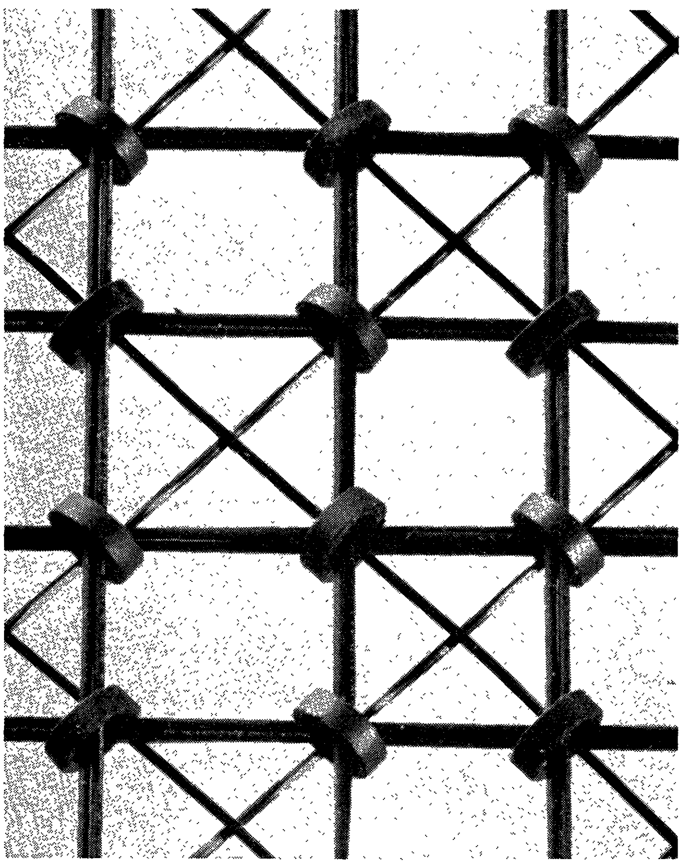


Fig. 8—Closeup of 32-by-32 plane.

tube on a bias level of about 30 volts. The control grid of the gate tube is pulsed with a 0.1-μsec pulse at the optimum moment so a "standard" Whirlwind pulse issues from the gate to indicate when a ONE is being read.

Layout and Packaging

A finished memory plane is shown in Fig. 7. The frame's outside dimensions are approximately 9.5 by 9.5 inches. All the windings consist of 32-gauge magnet

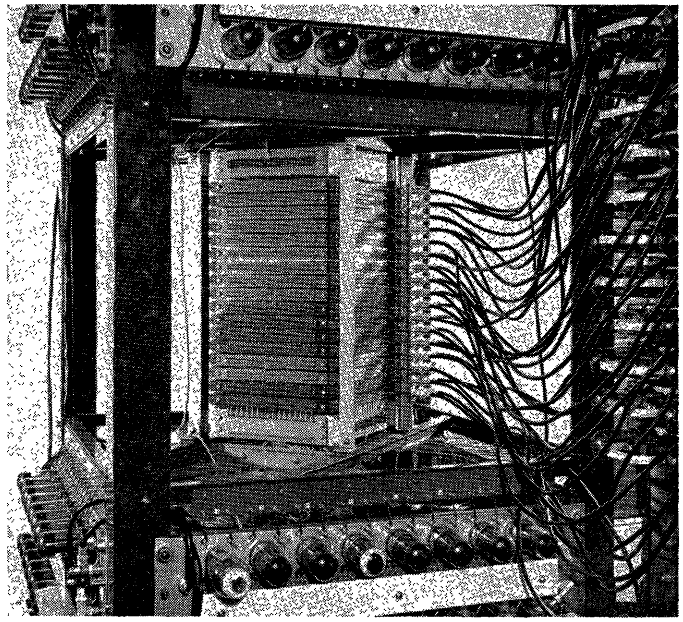


Fig. 9—Memory stack mounted in stall.

wire with quadruple-Formex insulation. Fig. 8 shows the cores and wires in some detail. The x and y pairs run vertically and horizontally, the sense winding runs along the diagonals, and the digit-plane winding runs horizontally (but is obscured in the shadow around the y pairs). Wiring time for one plane was about one man-week, including an intermediate test and final inspection. The intermediate test was performed when all the cores and x and y wires were in place but before the digit-plane and sense windings were installed; core replacement is relatively easy at this point. The test consisted of applying a sequence of current pulses to a given x line and observing the response of each of the 32 cores on that line by manually stepping the observing-scope probe from one y line to the next. This test was repeated for subsequent x lines, until the 1024 cores were completed. Cores which displayed abnormally high or low outputs were marked for replacement. About 1 core per plane was replaced.

The 17 finished planes were mounted in a stack or array, as shown in Fig. 9. Plane-to-plane connections are made by means of the vertical busses soldered into the slotted lugs. Digit-plane and sense-winding connections were made from the same corner of each plane to a mounting board of connectors for coaxial connection to another rack. Selection-plane-driving connections fan out horizontally at the top and bottom of the array. It takes 3 to 4 hours to replace either the entire array or any single plane.

Fig. 9 includes a view of part of the four-posted stall, or rack, in which the array is mounted. Fig. 10 shows a front view of the Memory Test Computer; at the left of the picture may be seen the memory stall. The Whirlwind I core memory is shown in Fig. 11. Selection-plane-driver panels are mounted on the four faces of each stall with tubes pointing outward. Visible in each stall above

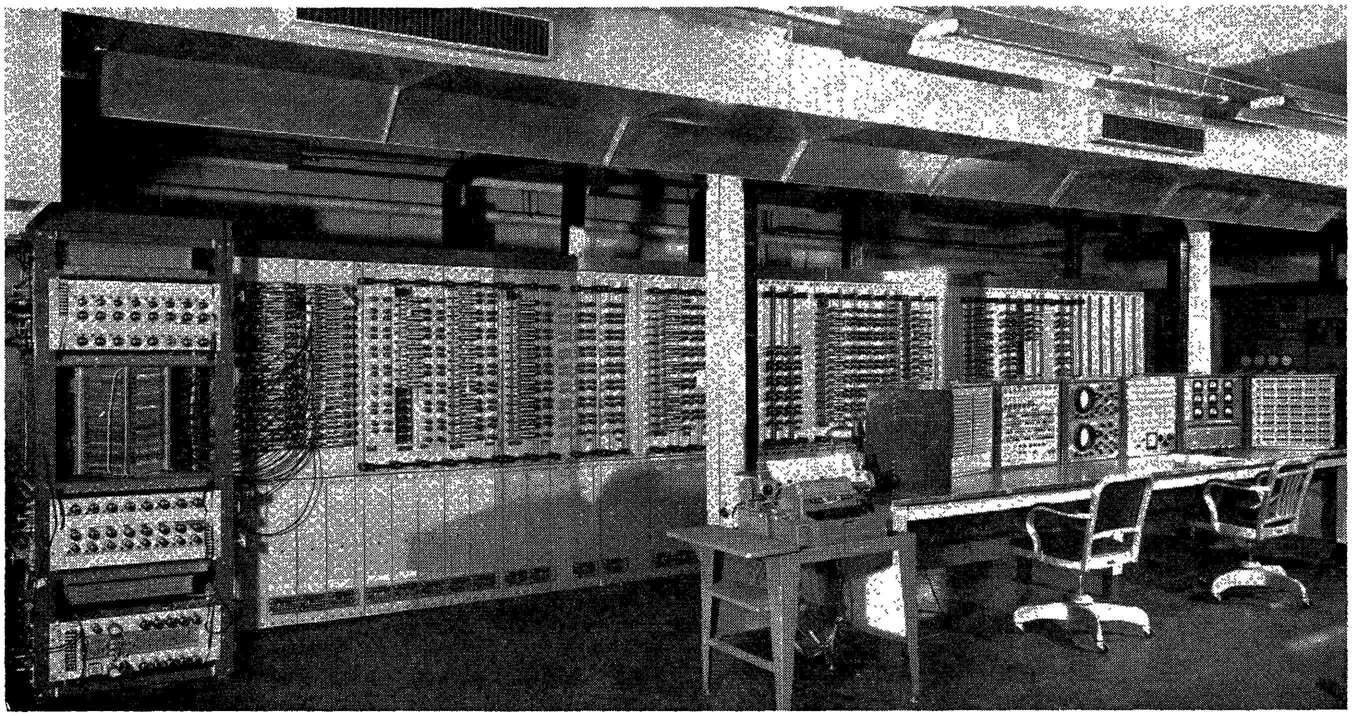


Fig. 10—Memory installation in MTC.

and below the selection-plane-driver panels are the two crystal-matrix switches. The general arrangement is such that temperature-sensitive components, such as cores and crystal diodes, are inside the stall, and large heat-dissipating components, such as tubes, are on the outside.

The sense amplifiers and digit-plane drivers are in plug-in chassis stacked in vertical racks next to the memory stalls.

Tests and Performance

Ultimate judgment on the reliability of this particular core memory must rest on its performance over the next year or two. Tentative evaluation may be made, however, from observations of performance during the 4 months that one bank operated in the Memory Test Computer and the 3 months of 2-bank operation in Whirlwind. In addition, much may be determined from the results of tests made on the memory to ascertain its tolerance to variations in the parameters significant to its operation.

Parameter Variations

Many conditions, or parameters, affect the operation of a core memory; driving currents (x, y, read, write, inhibit, and disturb), sense-amplifier gains, strobe time, ambient temperature, memory-information pattern, and repetition rate are good examples. These parameters are not all equally significant or equally easy to manipulate, and so some of them have, as yet, been examined in only a cursory manner. Because sense-amplifier gains have a simple, nearly linear, effect on operation they were adjusted and held at one setting during the tests. Ambient temperature is expected to be held within close

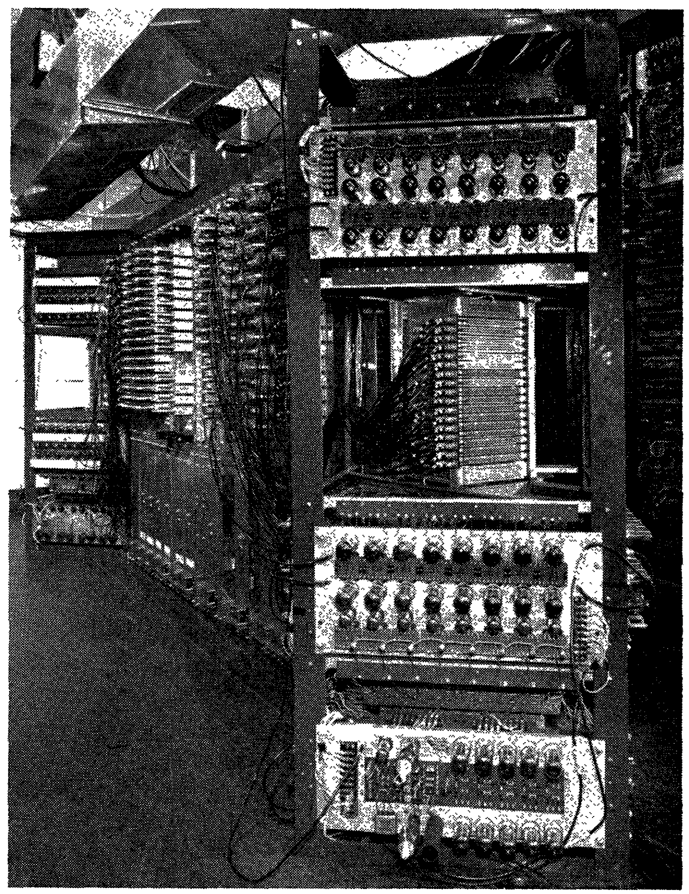


Fig. 11—Core memory in WWI.

tolerances in any operating machine, and a fair amount of information is available on the subject from the core-testing work; temperature was not controlled during the tests but recorded readings were kept. Memory-information pattern and repetition rate were controllable to some degree by the program being run; a program which seemed to give the most adverse pattern and rate was designed and used during most of the testing.

The tests were made on the Memory Test Computer, a high-speed, 16-digit, parallel machine of the Whirlwind type. The machine has a parity-checking system which computes whether each 16-digit word to be stored contains an odd or an even number of ONE's, stores the result of this "parity count" in the 17th digit, recomputes the count when the word is read out, and rings an alarm if the result does not check with the contents of the 17th digit. Although major reliance was placed on parity checking for detecting memory malfunction, there was also some programmed identity checking used.

The bias bounds of the sense gates' suppressor grids were chosen as a very convenient measure of the quality of the memory output. The upper bound (least bias) is the point at which errors occur because the gate is mistaking the largest ZERO output for a ONE, at the lower bound (most bias) errors occur because the gate mistakes the smallest ONE output for a ZERO. The bias difference, in volts, is a direct measure of the voltage difference at strobe time between the smallest ONE and the largest ZERO.

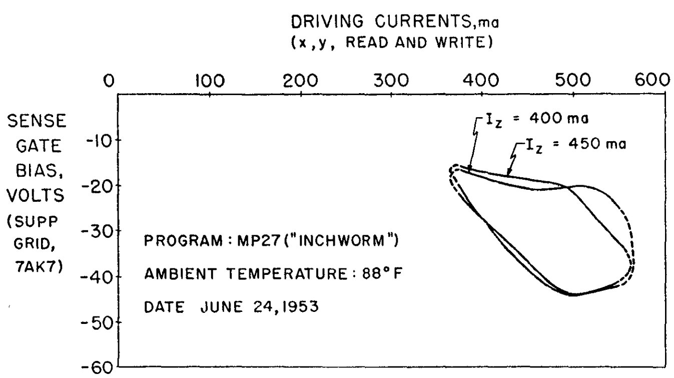


Fig. 12—Bias bounds versus drive currents.

Fig. 12 shows the bias bounds for all 17 sense gates as a function of the selection-plane driving-current amplitudes (x, y, read, and write). The program used was the so-called "inchworm" in which 16 words of instructions "bootstrap" themselves around the 1024 registers of the memory. The ambient temperature was recorded at approximately 88 degrees Fahrenheit, about 15 degrees higher than what is now believed to be optimum. Two curves are shown, one for digit-plane currents set at 400 ma and the other at 450 ma. The enclosed areas indicate how much the safe operating point of the memory bank may wander; recent circuit and adjustment improvements have enlarged these enclosed areas somewhat. Fig. 13 shows the bias bounds as a function of the timing of the strobe pulse. Time is measured from the instant the Read Flip-Flop is pulsed by the Start Read pulse. The three curves are for three values of selection-plane driving current, two extremes and one near optimum. A wide operating region is again indicated.

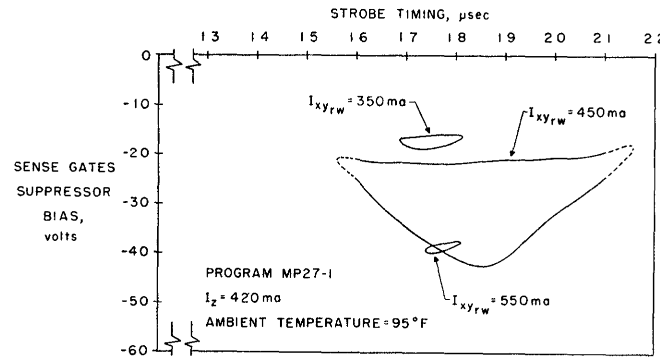


Fig. 13—Bias bounds versus strobe time.

Computer Operation

The first bank of memory has been in use in Whirlwind since mid-August, the second since September 5. There has been a steady improvement in their operation as the installation, which was an extremely hurried one, has been gradually cleaned up and made permanent and, also, as the process of debugging these relatively new equipments proceeded. The two banks have not quite been brought to an equal degree of reliability; this may be due, in part, to the fact that the cores in the first bank were not selected as carefully as those in the second so that output ONE/ZERO ratios are not as large. The demands on the Whirlwind computer are heavy, and only a few hours a month are available for further development work on its memory.

Parity alarms occurred, at first, about 3 or 4 times per week; at this writing (November 27) there has not been a parity alarm for four weeks. This comes to about 460 hours of useful operation or, assuming a 30-μsec average order time and 2 accesses per average order, it comes to slightly over 100 billion word accesses with each access parity checked and no error detected.

The exact nature of the errors which do occur is, as yet, not known. It is hoped that further work on the system will shed more light on the problem as well as reduce the error rate yet further.

Conclusion

The test results and experience obtained thus far on the two 32 by 32 by 17 banks of magnetic-core memory now operating as the internal memory of the Whirlwind machine indicate high promise for this type of storage. Reasonable engineering extrapolations of the results are being used in present work on a 64 by 64 by 17 bank which is expected to be in operation in the Memory Test Computer by January 1954.

Reliability Experience on the OARAC

ROBERT W. HOUSE†

Introduction

THE OARAC (Office of Air Research Automatic Computer) was delivered to the Aeronautical Research Laboratory at Wright Air Development Center in February 1953. After physical installation it required about two weeks for electrical installation and checking. Consequently, good computing time began the latter part of April.

During the month of May the machine was operated twenty-four hours per day, five days per week. Then operation was limited to sixteen hours per day until September when three-shift operation was resumed. There has never been more than one shift of engineering service. Consequently, if any trouble developed between 1630 and 0730, the machine remained inoperative until the next morning. This disregards such remedies as replacing blown fuses, etc. which the operator on duty might find obvious.

General Description

The OARAC is a general purpose machine with a large memory and medium operating speed. The number system used internally is the coded decimal system. The word length is eleven decimal digits, and the memory capacity is 10,000 words (magnetic drum) with an average access time of 8.5 milliseconds. Words are transferred throughout in decimal-serial, binary-parallel fashion. The input-output medium is 5/8 inch magnetic tape with a pulse density of 17 per inch and a tape speed of 50 inches per second. There are four information tracks and one clock track on the tape. The basic pulse repetition rate is 145 kc and average addition time is 90 microseconds exclusive of access time. Multiplication and division require about 8.5 milliseconds on the average, exclusive of access time. The over-all operating speed with optimum coding is about 100 operations per second with the present one-address system. Design has been completed to change OARAC to a two-address system.

The machine occupies about 68 square feet of floor space and is air conditioned. A transient-free supply of power is provided by a motor-generator set; the machine requires about 23 kw. It uses about 1,500 tubes and 7,000 germanium diodes.

The arithmetic and control circuits are composed of six basic plug-in units called turrets. They are "AND" circuits, "OR" circuits, gated blocking oscillators, trigger pairs, and voltage discriminators. There are about 1,000 of these turrets in the whole machine. In addition to these there is one "fuse" turret for each panel of fifty logical turrets.

† Wright Air Development Center, Wright-Patterson Air Force Base, Ohio.

The main cabinet is 15 feet long, $2\frac{1}{3}$ feet wide, and 7 feet high. Housed in it are all of the arithmetic and control circuits plus two tape mechanisms. The memory cabinet is $4\frac{1}{2}$ feet long, $3\frac{1}{2}$ feet wide, and 7 feet high. Housed in it are the drum, drum motor, and the record-playback matrix. Air with temperature and humidity controlled enters the bottom of both cabinets and is exhausted from the top by ducts to outside the building.

Reliability of Input and Output

As mentioned above the pulse density along the tape is 17 per inch with 5 tracks of information. Almost no difficulty has been encountered with head alignment or tape blemishes.

The eleven-digit word with the five-digit address associated with it are coded on tape with a leading guard code before the first digit and a following guard after the last. These two guards are actually wrong combinations of the number system used. If the leading guard is not present when the first clock pulse reads the code, an alarm is given indicating that the tape was not positioned correctly, or that the guard code for that word was missing, or that the alarms are failing. This check is used chiefly to position tape on mechanisms initially.

After the leading guard check is passed, the information that follows is checked for wrong combinations until the following guard is sensed. If a random error on tape produces a correct combination, this check will fail. Since errors usually occur in a consistent manner, this check has been fairly effective. This wrong combination check is also applied when words are transferred from the machine to tape.

Both input and output tape mechanisms have alarms for broken tape or exhausted tape supply.

Output is judged to be more reliable than input on the whole. One proposed change calls for an additional check. This check would add the first n words coming from tape and subtract the $(n+1)$st automatically. This means that when the tape is made the first n words would have to be summed to determine what the $(n+1)$st word should be. Ten seems like a reasonable value for n at present.

Reliability of Memory

Until the latter part of October the largest percentage of downtime was caused by the memory. Two construction errors were responsible for this. One was a bad solder connection for ground at the terminating end of the coax from the drum to the machine on six of the eight information lines. The other was the omission of a filter to half of the playback-record matrix. When these defects were removed, operation was greatly improved.

Information recorded on the drum is checked for wrong combination as it leaves the machine to be recorded and as it is received by the machine on playback. This check has been quite adequate to detect malfunctioning but is almost no help in localizing the defective component.

Reliability of Arithmetic Unit

The arithmetic unit has many checks at many places. The reason for applying some of them is obvious by their names, such as, sum exceed capacity, product exceed capacity, and divide by zero. In addition to these, wrong-combination checks are applied at two places in the chain of manipulations to achieve arithmetic results; three checks are applied (quite often throughout multiplication and division) to insure that the word used is a number and not an order, since numbers and orders are not in separate addresses on the drum (the difference is in the sign digit); and a check is constantly made to see that three drum revolutions in time do not elapse between end-operation signals (except on reading in from tape).

In addition to the types mentioned, there is one other alarm associated with a particular order. The order is designed to repeat a coding routine twice if a specified tolerance is not met. If the tolerance is not met the second time, this rollback order will call back to the start of the routine (preparatory to running a line of coding at a time) and then signal the alarm. This order enables an automatic repetition of parallel coding or an iterating program.

Parallel coding is considered necessary for lengthy computations. Some coding has not been done in parallel, and good results have been obtained; but usually the coding has been interspaced with routines that do use parallel checking. Consequently, if any components were bad the parallel checking would indicate it and cause their removal. The checks in the arithmetic unit are adequate for the most part to localize troubles rapidly, and only one additional check is being considered at present. This is the addition of one more wrong-combination check.

Reliability of Control Unit

The basic clock pulses for the machine come from one fixed track on the memory. In addition there is one other track with one playback pulse per revolution. By means of this single pulse the counting in the two basic machine counters is checked with each drum revolution. If either is out of synchronism, an alarm is given.

In the present address system five numbers specify an address. The first two designate the track and run from 00 to 49; the next three designate the position around the track and run from 000 to 199. Because of this it is possible to designate an address which is nonexistent, e.g. 48307, the three being wrong. There are checks at two points in the control circuits to give an alarm for a nonexistent address.

On the whole the checks in the control unit are quite adequate to detect malfunctioning, and to a large extent they are helpful in localizing the trouble. One additional check will be put in when the conversion to the two-address system is made. This will be a wrong-combination check on addresses.

Reliability of Power Supplies

When the machine was originally constructed, five power supplies were subcontracted for and the rest were made by the General Electric Company. No appreciable trouble has been given by those supplies made by G. E., but a large amount of downtime has been caused by the other supplies. The difficulty was realized before the machine was delivered and an agreement was made by G. E. to replace the faulty supplies. These new supplies were received and installed in October.

Reliability of Components

When the machine was originally designed, the 1N52 germanium diode was the type designed to be used. However, as checking of the various parts of the computer progressed the number of diodes failing became prohibitive. At that time there was no temperature or humidity control on the air used to cool the machine. Air was merely blown through the machine at high velocity. The decision was made to change over to the 1N63 diode. At about the same time a five-ton air conditioner was installed to supply part of the air used for cooling, and the back voltage on the diodes was decreased by five volts. It was not ascertained which, if any, of these actions was effective, but diode failure dropped to what was considered a reasonable rate. An estimate of the present diode failure rate is ten per week (24 hour operation—5 days). Over 90 per cent of these fail test specifications on the bench and are not in-service failures.

Two types of tubes are used extensively in the main machine. They are the 12BH7 and 2C51. The estimated life of the 12BH7's is 12,000 hours. The 2C51's, which are not as numerous as the 12BH7's, are considered to have a much longer tube life. It might be of interest to note that one brand of 12BH7 has been totally unsatisfactory in the trigger pairs in the main machine, the trouble being instability.

Conclusions

It is felt that the machine has not been in operation long enough to make any conclusive estimations of how good it is. Most of the personnel working on the machine, including mathematicians and operators, had very little, if any, knowledge about the machine before it was delivered. Because of this, as personnel become better trained, the operating efficiency should increase.

Discussion

J. C. Cochran (U. S. Department of Defense): How have the sockets of the Alden plug-in units held up?

Mr. House: I have no complaints. They work out fairly well. We have had on occasion some trouble with pins in the base being shorted, but only very rarely, and these can be corrected quickly by merely drilling a hole through the base.

C. W. Watt (Massachusetts Institute of Technology): What air temperature and humidity are used in the air input?

Mr. House: This is completely controlled by the conditions of the room; we keep the control set at 71 degrees and 50 per cent relative humidity.

C. T. Schaedel (Consolidated Vultee Aircraft Corp., Fort Worth Div.) Please explain further your comment concerning technicians and diode failure due to improper handling. You used the words "kill the diode."

Mr. House: If a technician is trouble-shooting a defective turret on one of these plug-in units, it is quite possible that he may short out something and put an excessive voltage across the diode. On these occasions the diodes are of no use afterwards.

M. W. Marcovitz (Burroughs Adding Machine Corp.): What type head is used?

Mr. House: It is made by G. E. especially for this computer.

Mr. Marcovitz: What is the pulse density on the drum?

Mr. House: It is about forty to the inch.

Mr. Marcovitz: Is there a read and write amplifier for each head?

Mr. House: There is not. In fact, there are only four amplifiers for the four binary bits associated with each decimal digit.

B. Aaront (R.C.A.): What is your criterion for determining good time?

Mr. House: By good time I mean total on time minus the time used to repair troubles and the time used for preventive maintenance. In good time I would include the time for checking out a problem.

Operating Experience with the Los Alamos 701

WILLARD G. BOURICIUS†

THE NEED for large scale calculations has grown exponentially with time at Los Alamos. In the fall of 1950 we had one Card Programmed Calculator, we now have six. About a year and a half ago, the Los Alamos Maniac was put into operation. And still we needed more calculations, so a 701 calculator was ordered to help fill that need. Our 701 was delivered last April and placed under the administrative control of a group called T-1, within the Theoretical Division.

Before telling you the performance statistics on the 701, I think it a good idea to give you a description of how the 701 is managed, including the kind of people who program and code for it, and the advantages they gain by the present operating system.

We operate our 701 on a self-service basis. This method of operation means that anyone having a 701 problem may program, dode, debug, and run his own problems with only as much aid and assistance from Group T-1 as he requests. The self-service system for the 701 is an outgrowth of our experience with Card Programmed Calculators. Originally, the Card Programmed Calculators were available without operators so that the scientists had to run their own problems. This turns out to have several advantages: They avoid communication delays and difficulties because there is no intermediary between them and the machine; they keep in intimate contact with their problems; they get to know the full capabilities of the calculating machine and very quickly make use of this knowledge; and they obtain their answers sooner. A further advantage to the self-service system is that the responsibility for the problems is placed where it rightfully belongs, on the problem originators. In actual practice, the majority of our high-caliber physicists and mathematicians code and debug their own problems.

In order for 701 users to reap all of these advantages of the self-service system, Group T-1 has assumed several responsibilities. I will discuss these in detail.

In the first place, we have provided appropriate educational facilities to potential 701 users. We have furnished manuals on 701 coding, taught several classes on 701 coding and 701 operating, and arranged for an "Advanced Coding Seminar." We have encouraged attendance at all these classes and the seminar and have done everything we could do, within reason, to disseminate knowledge about coding for and operating of the 701.

We have also coded and debugged a comprehensive set of utility programs and a library of subroutines of commonly used mathematical functions. We have placed these programs in convenient locations and publicized their existence. These utility programs cover every phase of communication between various input-output units of the 701. For example: The card-read programs will read decimal, octal, and binary cards, fixed point and floating point constants, absolute, regional decimal, or regional binary programs. The term "regional" will be explained a little later. The print programs will print half or whole words, decimal or octal, absolute or regional. There are several programs which will transfer information back and forth between the drums or tapes and the electrostatic storage.

We have standardized and encouraged the use of a relative coding system called Regional Coding, which was devised in its original form by the Applied Science Department of the International Business Machines

† Los Alamos Scientific Laboratory, Los Alamos, N. M.

Corporation. Regional coding is a method of coding wherein every address is prefixed by an identifying combination of two digits and a letter. A regional code is coded as if it were to be located at the beginning region of the electrostatic storage, but during the assembly stage this regional code is relocated into any specified part of the electrostatic storage.

All regional coding at Los Alamos is done in the decimal system. Large problems are sub-divided into small units, and each unit is coded separately. These units are brought together and assembled into a complete code by the 701 itself. The utility program which accomplishes this assembling reads the regional code and prints on a listing both the original regional decimal code and the relocated octal absolute code. For every sheet of this listing, and in a one-to-one correspondence, a binary punched card is produced which contains the code in absolute binary, together with a check sum. This deck of binary cards contains the final code in a condensed form of forty-four instructions per card. These instructions may be quickly loaded into the electrostatic storage at the rate of sixty-six hundred instructions per minute. You will appreciate how convenient this speed is when I tell you that it is not uncommon to load several hundred different programs into the 701 during the course of a day's operation. When a code is in regional form, insertions and deletions in it are made quickly and accurately, without leaving the logic in a snarl. One further advantage of regional coding is the fact that one can study the listing, which was printed during the assembly of the absolute code, for logical errors before actually trying out the code on the calculating machine.

Every morning, time is allotted to the individuals requesting it. Between twenty and thirty different people debug or run their problems during a twenty-four hour day. There is a dispatcher sitting at a desk inside the 701 room who sees to it that the correct individual gets his or her assigned time. This dispatcher also keeps the 701 log, writing down the disposition of the 701 time to the nearest minute. The information in this log is punched onto cards and tabulated, to give monthly statistics on 701 performance.

When a 701 user reaches the debugging stage of his problem, he will find several utility programs which will be of great value to him. There is, for example, a "print memory" utility deck of binary punched cards, which will print out the contents and the location of every half-word in the electrostatic storage memory, skipping all storage locations that contain zero. There are tracing utility decks which trace the progress of any specified part of the code in a dynamic manner, printing out, in both the octal and decimal systems, the contents of the accumulator and the multiplier-quotient registers at the end of every operation.

All of the utility programs and function subroutines are available immediately to the 701 user in both regional and absolute decks. Furthermore, we have developed "regional binary" cards containing forty-three instructions per card so that these utility programs can be read into any region within the electrostatic storage that happens to be available. Most of these programs are controllable by punched cards. The whole scheme is devised in such a way that only three buttons on the console are ever pressed; the "reset and clear memory" button, the "load" button, and the "start" button. Although any novice can learn to operate the 701 console in about twenty minutes, our most adept operators leave it strictly alone. They have found it preferable to put information into the 701 by means of punched cards than to fiddle with the keys on the console. By using punched cards, fewer errors are made, a record of what went into the machine exists, and, in general, machine time is saved. It is only the rapid communication between the 701 operator and the 701, provided by the card reader and the line printer, that makes possible these debugging techniques. Most of the errors are pinpointed by people sitting at their desks poring over a listing from the printer.

Many exploratory problems cannot be scaled in any reasonable length of time, and, for that reason, floating point subroutines have been provided. Also, we have available two decimal interpretive systems which have proved useful for such problems. The most used of these systems is called Dual Coding and consists of a floating point abstraction of the 701 commands. Dual got its name from the great ease with which a coder can switch back and forth between fixed and floating point orders in the same problem. Another great advantage to coding problem in Dual is that one can code in Regional and thereby make use of all of our utility programs. The other interpretive scheme is called Short Hand Coding, commonly called SHACO, and is a three address ten significant figure floating point decimal system. A Dual problem is on the 701 about thirty per cent of the time, and a SHACO problem is on about five per cent of the time.

TABLE I

Table of Percentage of 701 Time According to Usage

	April	May	June	July	Aug.	Sept.	Oct.
Good Calculate Time	63.5	72.1	78.2	66.2	84.0	82.2	77.9
Maintenance Time	30.2	22.8	16.1	24.5	10.5	14.5	17.8
Lost Time—701 Error	4.6	3.2	3.2	6.2	2.3	1.4	2.5
Lost Time—Human Error	1.7	1.9	2.5	3.1	3.2	1.9	1.8

Table I gives the 701 performance statistics. The Lost Time due to human errors includes the time lost because the operator made an error such as having the wrong print panel in the printer or putting the cards into the card feed in the wrong order, and other things of a similar nature. The Lost Time due to the 701 is the time charged to an error which has been traced, either at the time of occurrence or later, to the 701 malfunctioning. The Maintenance Time is the time the IBM engineers use for both scheduled and unscheduled maintenance. Of this maintenance time, eighty to ninety percent is scheduled. The Good Calculate Time is the time remaining when the maintenance and the lost time are subtracted from the whole. As you can see by the figures, the lost time due to 701 errors and human errors, together, amounts to only three or four per cent. Maintenance Time seems to be leveling off to about fifteen per cent, and the Good Calculate Time seems to remain about eighty per cent. We began operating twenty-four hours a day in August, but I hestitate to draw any conclusions regarding the effect this has had on the per cent of Good Calculate Time.

Let me summarize the following results of our decision to employ the self-service system for operating our large calculating machines. In January, 1952, no one at Los Alamos knew how to code for the 701. By April, five people knew enough about coding to write SHACO. By November, ten people knew how to code fairly well, but not how to debug the code. By the time our 701 arrived on the last day of March, 1953, about twenty people knew how to code very well and how to debug fairly well. At the present time, about seventy-five people at Los Alamos are coding and debugging 701 problems, and by next February, when our second 701 is in operation, I believe their number will double.

In conclusion, I should mention the host of people who contributed to the success of our 701 operation. The work was done under the general supervision of Bengt Carlson, the leader of group T-1. The Short Hand Coding system and most of the original utility programs were developed by William Anderson, Allan Benson, Ruth Freshour, Donald Monk, Dorothy Monk, Burton Wendroff, Eugene Willbanks, Edward Voorhees and myself. The Dual Coding system was developed by Bengt Carlson and Stewart Schlesinger. Most of the up-to-date utility programs, such as the assembly program, were coded by Ivan Cherry, Margaret Fellows, Paul Harper, Clifford Moss, and Dura Sweeney. Throughout, we have had the complete co-operation and advice of Lloyd Hubbard and other members of the IBM Corporation.

Discussion

J. W. Leas (RCA Victor): During the period of good calculating time, was the 701 giving useful output continuously, or were there periods of inoperation; that is, what was the useful work percentage?

Dr. Bouricius: The slide that was shown was on the basis of 24 hours a day, five days a week, and over the period of the last three months, there were no periods of inoperation except over the weekend.

J. Rosenberg (GE): Do you have a breakdown on the causes of error or breakdown on the 701?

Dr. Bouricius: The 701 in my opinion is a pretty well-balanced machine. By that I mean that if any one component or any one part were improved by say a factor of 10, it wouldn't increase too much the reliability of the machine as a whole. Nonetheless I feel that the weakest part of the machine—at least the part that is most variable—is the electrostatic memory, and I would judge that roughly half of the errors and half of the maintenance time are due to it.

J. Jay Wolfe (George Washington University): What percentage of the good calculating time is spent in debugging?

Dr. Bouricius: I prefer not to answer that question without some explanation because it depends very heavily on the type of problem you have. On some types of problems, you spend more time debugging than you do running them, and then there are other problems that turn into parameter studies in which you spend much more time operating than you do in debugging. I think that the problem isn't too pertinent except as a matter of manpower. If you have a lot of problems that need debugging—research type problems—you have to have more manpower. Actually, when we started we spent about two-thirds of the time debugging. Now we are down to about half-time debuzzing. Maybe later on it will be down to one-third, but maybe not.

Clyde E. Willis (Goodyear Aircraft Corp.): Do you also operate the CPC's on a self-service basis?

Dr. Bouricius: Yes. We didn't do this deliberately as a matter of policy. It just grew that way. We carried it over from the CPC's to the 701.

Mr. Willis (Goodyear Aircraft Corp.): Are there not some disadvantages to a self-service basis?

Dr. Bouricius: Yes, there are. One of the disadvantages is that it requires more tactful people. You see, we operate the 701 just as you would operate a Safe-Way store; we manage it. It's self-service. We treat the people who come in to use it as customers and in this connection we try to reason that the customer is always right. Now, the advantages are that the good people learn the machines. They learn them very thoroughly and they squeeze every ounce of capability out of the machines. The disadvantage is that there seems to be no defense against stupidity. We have to adopt a very charitable attitude. When people first come in they are very hesitant; they don't know exactly what to expect. But if we treat them like long lost brothers they soon learn self-confidence and they are not adverse to asking questions. The communication between them and us is rapidly established and if we don't make them look foolish when they do ask a stupid question, pretty soon their questions become much less stupid and then in a very few days or weeks their questions become rather sophisticated.

Dorothy T. Blum (Department of Defence): Do you believe that your self-service system would work effectively in a *large* computer installation which handles diverse problems?

Dr. Bouricius: Self-service systems have grown up in other places than Los Alamos. Northrop Aircraft Corporation has had it, I believe, for some years. North American in Los Angeles is establishing the self-service system with their 701, and all their engineers are going to code their own problems and run them on the 701. I see no reason why this would not work for a *large* computing organization.

Joseph V. Natrella (U. S. Air Force): Could you clarify the good time figure of about 87 per cent for your 701? Was scheduled maintenance included in the good hours? Is 701 usable 85–87 per cent of total time scheduled for production?

Dr. Bouricius: The columns of figures in Table I add up to 100 per cent. Since August this has been 100 per cent of a 24-hour day, 5-day week. This means that scheduled maintenance was not included in the "Good Calculate Time."

Natalie Coplan (U. S. Air Force): I would like a description of the code which assembles "regionally" coded routines into one code.

Dr. Bouricius: Our assembly program will assemble programs that have up to two hundred different indices. It reads the cards, relocates, prints, and punches in binary cards at the rate of seventy instructions per minute. Later on, for changing the assembled code by insertions or deletions, we have a program that will accomplish this at a speed of about seven hundred instructions per minute. A complete description, and the code itself, for any of our programs, will be sent to anyone who writes for it.

Acceptance Test for Raytheon Hurricane Computer

F. J. MURRAY†

Introduction

THE HURRICANE digital computer is a large-scale digital computer developed for the United States Navy by the Raytheon Manufacturing Company and installed at the Naval Air Missile Test Center at Point Mugu in the spring of 1953. It has many novel features including highly developed input and output equipment, magnetic-tape storage with an optical locating device, and a checking system which was designed not merely to detect errors but to prevent loss of any valid information available to the computer.

While the primary purpose of an acceptance test is as usual to determine whether specifications have been met, in a novel device of this type incorporating such unusual features, the acceptance test also serves as a preliminary evaluation of these features. Our testing procedure included a step-by-step test of each logical claim made for the machine as well as a number of over-all operation tests. For a general purpose computer this testing procedure is more rigorous than normal engineering testing would be in most respects. The exception is in testing under marginal operating conditions. This was not permissible in the present circumstances.

The magnetic-tape storage was extremely impressive. The checking circuits not only kept the computations free of error but because of their diagnostic effect, the machine itself was brought to a remarkable state of operational perfection. Thus a system difficulty which would occur about one time in a million repetitions of one specific order in one set of circumstances was detected and eliminated.

In the present paper we describe the acceptance tests. We append a detailed description of the computer itself.

The Test Phases

There were two objectives of the acceptance tests. One was to demonstrate that the various parts of the computer were capable of functioning as claimed, and the second objective was to establish its ability both to process a large amount of information and to act as a large-scale, general purpose computer by solving a system of simultaneous differential equations.

These tests were divided into phases. Phase I is concerned with demonstrating the ability of the machine and was subdivided in three parts. Phase I_0 refers to those parts of the test which could not be demonstrated by means of a regular programmed procedure—for instance, the demonstration of the adding abilities of the checking circuits. Phase I_1 is a regular programmed procedure demonstrating the ability of the machine to operate under approximately normal conditions for a fixed period of time, usually five minutes. In Phase I_1 tests, the checking circuits were relied upon to insure the correctness of the operation. In the Phase I_2 test, the individual functions of the machine were established with printed evidence. Furthermore, these Phase I tests were subdivided according to the units of the computer as given in the appended description.

Phase II consisted of the evaluation of a rational function with numerator and denominator of the *4th* and *3rd* degrees respectively, for 17,728 pairs of values of two variables x, y. This was done in five minutes and ten seconds. It was intended to demonstrate the ability of the machine to process a large amount of information. Phase III was the solution of a large system of differential equations corresponding approximately to a description of the flight of a guided missile for 65 seconds. The solution of this problem was already available.

As a computer or device for handling information, the machine has certain specified operations to perform. These include the arithmetical operations and the transfer of information between different media, for instance from magnetic tape to acoustic line storage. For many of these operations it is possible to obtain a permanent printed record. This constitutes the Phase I_2 tests. However, normal operation in general will not involve production of this printed record and will proceed faster. To demonstrate the action of the machine under normal operation, the Phase I_1 tests are used, and their correctness is dependent upon the check circuits and coded checks.

However, in the operation of the machine, there are two processes which normally produce no external evidence. These are the checking processes and the selection of the operands by the central control. Since the memory is serial, it is very desirable to save time by using the first operand which becomes available, rather than a specified operand in the half order being used. This is a function of the central control and normally produces no external evidence. However, proper coding can utilize this function to produce variations of the time of the machine and this was demonstrated as part of the I_0 tests for the central control.

Checking Processes

Five checking processes are used in the machine.

1. *Transfer weight count checking.* Associated with each word consisting of 30 binary bits and sign, is a transfer weight count defined as follows. The given word

† Columbia University, New York, N. Y.

can be expressed as plus or minus times the octal expression for this number. Now add the ten octal digits in this expression and add one if the word has a plus sign and add two if the word has a minus sign. The transfer weight count is a number between 0 and 15 congruent to the result just obtained, modulo 16. At each transfer of a number in the machine, the transfer weight count is recalculated from the transmitted word and compared with the transfer weight count transmitted with the word. A discrepancy will cause the machine to stop.

To test this function the counter used in recalculating the weight count must be shown to be able to function as an adder modulo 16. There will be four binary digits in the result and one must show that a discrepancy in any individual place will cause the machine to stop. In some cases the transmitted transfer weight count is entered into the counter in complemented form so that the final result is in a prescribed form. In these cases the ability to enter the complemented form was tested. Note that the fact that a false transmitted weight count would stop the machine does not prove that the checking procedure is working correctly, since any addition error except possibly one could also yield an alarm.

To check the adding properties of the transfer weight counters, words were constructed which would enter every possible octal digit into every possible configuration of the counter. The result in the counter could, in practically every case, be read from certain lights on the machine.

Normally the transfer weight count is obtained serially, digit by digit. In some exceptional cases, the transfer weight count is obtained by a parallel addition of all the digits at once by a "tree adder." This occurs in the relay storage associated with the printers. Theoretically this requires that the sum up to each place take on all possible values mod. 16 and to each following place all possibility should be assigned. However this involves normally too many possibilities and a sampling process is necessary.

2. *Arithmetical counts.* Each operation of the arithmetical unit is also checked by performing a subsidiary computation on the quantities x_c obtained from the operands x as follows. Express x as a radix 32 number. Add the digits in this representation and call the result x_c. For each of the operations of the arithmetical unit, it is possible to find an appropriate check formula on the quantities x_c. For instance, $x+y=z$ has corresponding the check formula $(x_c+y_c+\bar{z}_c)_c=31$ where $\bar{z}$ denotes the complement of z as a binary number.

In the arithmetical checks two arithmetical operations are used, i.e. addition and multiplication, and for shifting operations a certain shift in the "primary transfer weight counter." The arithmetical count register was tested as an adder as in the above. Multiplication is based on the use of eight serial adders and these also were tested. The above mentioned shifts and the compounding of operations as needed in the various checking formulas were also demonstrated.

3. and 4. *Operand addresses and operational code matrix tags.* The address of an operand as given in a half order has two parts, a line tag to identify the mercury line in which it occurs and a word tag to specify the time at which the operand is available. When a line is used, it produces a "tag" which is compared with the line tag in the operand as a check. Similarly, the choice of a word in the line produces a word tag which can be used as a check. The operational code as part of a half order is used to select a matrix to perform the specified operation. The operation matrix also produces a "tag" which is compared with the operation code in the half order which is being executed.

The proper functioning of these last "tag" checks in 3 and 4 is somewhat difficult to check since a deliberately introduced error may stop the machine in many ways. This was circumvented as follows: Suppose a certain error will lead to an alarm under a specified condition, for instance a discrepancy between a specified bit in the tag and the corresponding bit in the original with which the tag is compared. Now one makes these bits equal and shows that the machine proceeds without the alarm. Now one of these bits and nothing else is changed and the alarm results. One then disables the alarm. The machine should proceed with the discrepancy as if nothing were wrong. This shows that only the specified fault stopped the machine.

The 5*th* type is the vane checks used to check the printer output. This is a "tag" check in the sense that a signal is returned from the printer and is subject to analogous difficulties. The tag is produced by the five teletype bars, each of which has two positions. Each bar was held in a specified position and permitted to perform operations consistent with that position. Then an operation was performed inconsistent with this position and an alarm obtained.

The Phase I_2 Tests

In the regular programmed portions of the phase I test, and in particular phase I_2, it was not possible to demonstrate every combination of operands. However, the operations of the arithmetical unit can be analyzed into basic operations of addition of positive numbers, complementation, shifting in registers, transferring, certain remote carry possibilities such as occur in multiple precision operations, sign operations, and the combination of these basic operations into such operations as algebraic addition, subtraction, multiplication and division, normal shifts, square-root shifts, and transfers. In the I_2 tests it is supposed that the basic operations can be tested individually and independently of each other and that the possibilities for combining basic operations into arithmetical operations can be tested independently of what is happening in the individual basic operation. Thus if basic operation a has m possibilities and basic operation b has n possibilities and these are to be combined to yield arithmetical operation A, one might require that operation A be tested for mn possibilities. However, our independence assumption would mean that we would test the m possibilities for a and the n for b once and for all and only one combination of these to yield A. This was not carried to this extreme but one was forced in this direction by the multiplicity of the possibilities.

One has to analyze even further the basic operation of the addition of positive binary numbers into what can happen at the individual binary places. One assumes that the actions which occur at two different binary places are independent unless there is a carry relation between them. Thus the no carry possibilities for the thirty binary places can be tested by three additions, i.e. of the number zero (which we write 0^{30})[1] to itself, and of 0^{30} to 1^{30}, and of 1^{30} to 0^{30}. If the independence assumption were not made, three to the thirtieth power tests would be needed. Of course the carry possibilities have to be analyzed into cases according to the number of places involved. However, places associated in a carry operation are adjacent and the sets can be tested independently. Thus if two places are involved, we must have 01 added to 01. This can be tested by adding $(01)^{15}$ to $(01)^{15}$ and $0(01)^{14}0$ to $0(01)^{14}0$ avoiding the overflow. For three places, there are three carry possibilities and three shifts are to be applied to each possibility yielding nine additions. However, for each increase in the number of places the carry possibilities are tripled and the number of shifts increased by one and so after three places certain characteristic long carry possibilities were tried and this was considered sufficient.

Test of Specific Claims

In addition to the general claims made for the machine there were a number of special claims made for specific components. Thus each table (input desk) had a monitor feature which consists of a tape reader, such that the tape prepared by the operator is monitored to be identical with the tape read. This was tested in detail. The magnetic-tape reader had a number of features which were intended to prevent the accidental destruction of valid information, either by illicit order combinations or by misoperation. These were also tested.

Each table was given a regular performance test under normal conditions involving its full range of operation. In addition the monitor feature was tested by trying to press every key but the correct one.

An accidental write order to a tape unit can destroy valid information. A "write" order is preceded by a "hunt prepare to write" order, and similarly with the "read", and a mis-sequence of operation stops the machine.

[1] Here we use the following notation for brevity: The expression x^y means that the quantity x is to be written down y times in succession.

Much of the effectiveness of the machine depends on the buffer storage facility associated with the tape units. These permit the tape units to function independently of the remainder of the machine. A detailed test of this functioning was made.

Appendix

Description of the Hurricane Computer

The Hurricane Computer is an automatic electronic sequence calculator, with an acoustic delay line fast memory of 1,152 words, and an external magnetic-tape storage of four units each with 100,000 word storage. The computing word has 30 binary places and sign. The machine uses the binary system. Information is transmitted in the machine in serial form but the arithmetical unit is based on parallel registers. The basic pulse rate is 3.77 mc.

The machine operates on a four address code. Two words constitute an order. The first word or first half order contains the addresses of the two operands and the operational code in binary form for the operation to be formed. The second half order contains the address to which the result of the operation is to be sent and the address of the first half order of the next order. After performing an order, the machine proceeds to take in the next order to its central control. To do this the two half orders are stored in sequence in the memory.

The basic operations consist of addition, subtraction, multiplication, and division, shifting, branching, and extraction. The operations for which codes are available are constructed from these; for instance one has addition and subtraction, multiplication with round off, and division with round off. Transfer is essentially the addition of zero to the given number and placing the result in the new location. Transfer can occur with a change to absolute value or to negative absolute value. Orders may be modified by an ordinary addition operation or by a special substitution operation. The results of computation may be used to control the sequence of computation by means of a branch order. The second half of a branch order contains two possible addresses for the next order. The choice of these possible addresses depends on either an inequality or the existence of an equality depending upon which of the two branch orders is used. Extraction is a process of producing a new word by considering only certain digits of a given word, the other digits in the result being zero. For multiple precision, there are orders which produce the high order part or lower order part of the product of two words. Double precision addition and subtraction are obtained by two orders. There is a controlled shift, a normal shift in which the number ends in a position with the first two bits 01, and a square root shift. There are compound operations which permit the easy programming of the floating point operations of addition, subtraction and multiplication. There are three operations associated with the external memory.

The speed of operation has been described as follows:

Access time to internal storage (average)	125 μsec
Addition time	48 μsec
Multiplication time	285 μsec
Division time	430 μsec

The machine requires a power of 37 kw. It has 5,200 tubes and 18,000 germanium diodes.

The machine is based on the combination of a partly serial memory and parallel registers in the arithmetical unit. The memory consists of 36 mercury acoustic delay lines, 4 of which are associated with external memory. Each delay line contains 32 words. The content of each line appears as a sequence of 1,024 pulses or pulse omissions corresponding to the bits of these 32 words and an individual word is available only during a one thirty-secondth of the time. In order to minimize the delay caused by this, the central control operates under a complex set of rules in selecting the first of the two operands in a half order.

The Hurricane Computer is designed to be self-checking to the point that any single error will stop the machine. Accompanying each word is a transfer weight count, consisting of one plus the sum of the octal digits modulo 16. At each transmission of a word from one place to another in the machine this transfer weight count is reformed from the transmitted result and compared with the transmitted count. A discrepancy will cause the machine to stop with an indication of the cause of the difficulty. The arithmetical operations are checked by forming sums of the digits in the words when the word is expressed with 32 as a radix. The operand addresses in an order and the operational codes are back checked when these are used in the central control. When information is printed out on either the output printer or the directly connected printer, a "vane check" is made which shows that the proper type has been used in the teletyper. There are a variety of checks on the external memory units which warn of improper action and especially improper action which can destroy valid information on the magnetic tape. Physically the computer is divided into units.

The Problem Preparation Unit

This consists of two tables which are duplicates and a tape converter unit. One operator may use either table to enter the desired information onto a punched paper teletype tape. This prepared tape is inserted into the second table and a second operator repeats the process used by the first operator and consequently prepares a second tape. While this second tape is being prepared, it is compared by a reading mechanism on the table character by character with the first tape, and a discrepancy will prevent the second operator from proceeding. On both tables, a transfer weight count is formed for each word. After the second tape has been prepared it is used in the tape converter to produce a magnetic tape, suitable for insertion in the external memory unit. The reading of the paper tape is transfer weight counted.

The Output Printer

This consists of two parts. One of these reads a magnetic tape and converts the information on it into relay storage from which it is printed on a teletype printer. The transfer from tape to relay storage is checked by comparing the transmitted weight count with the weight count formed in the relay storage by means of a "tree circuit." The choice of print bar in the teletype is compared with the relay storage by means of a "vane check." The output printer is capable of printing the content of a magnetic tape in a format consisting of parallel columns. This format can be controlled either by controls mounted on the magnetic-tape reading unit or by signals on the magnetic tape itself.

The External Memory Unit

This unit consists of five racks, four of which contain identical tape reading and recording units and a fifth rack which controls the action of the other four units. The tape used has its information in magnetic form on one side in a block arrangement in which six bits appear in parallel. Seven magnetic heads are used to read this information simultaneously with an additional line of synchronizing magnetic marks. Six rows of six bits each correspond to a word and contain the sign, the thirty bits of the word and the four bits of the transfer weight count. Each block contains 192 rows or 32 words. On the other side of the tape, the block number of the block appears in black marks in binary form. These black marks are read by a photo-cell and the hunting process is based on these.

The four instructions for the external memory unit are hunt read, hunt write, read and write. In the hunting operations the tape is positioned by a servo mechanism until the block under the reading head has the proper value. The hunting process consists of transferring to the "hunt register" the number of the desired block. When the hunt operation is given the tape moves to place the next highest block under the head and the block number is read optically. If the block number read is higher than that desired, the tape-moving mechanism reverses. The tape moves until the block number read agrees with the desired number in the hunt register. Any further reversal of direction is a mis-operation and causes an alarm. If the external-memory unit receives another signal from central control before an operation has been completed this new signal is stored until the unit can act in accordance with it.

Each of the four units contains two mercury acoustic storage lines or registers of 32-word storage capacity. The process of passing from magnetic tape to acoustic line storage involves reading the six bits of the row on the magnetic tape, simultaneously storing these on flip flops, converting to serial form and then entering the lower acoustic delay line. The contents of lower register may be transferred to upper register and vice versa. The upper line in each unit is used like a mercury line of the internal memory. The passage from magnetic tape to acoustic storage is transfer weight counted. In the reverse process, the signal in the magnetic head at the time of writing is transfer weight counted but not the pulses on the tape. The information on the tape is always handled in blocks.

The tape characteristics have been described as follows:

size	$\frac{1}{2}$ in. × 0.0029 in. × 900 ft.
no. of words in each block	32
no. of blocks to each tape	3,150
no. of channels	6 information, 1 synchronization
reading and recording rate	300 words/second
pulse density (along length)	72 pulses/inch
tape speed—forward	30 inches/second
tape speed—reverse	60 inches/second
start-stop time	4 milliseconds
start-stop space	$\frac{1}{4}$ inch

The Clock

This consists of two parts. One part contains the fundamental 3.77-mc oscillator and a delay line containing a single pulse, which appears at the output in intervals of 1,152 fundamental cycles and marks the start of the major cycles. Each major cycle contains thirty-two minor cycles and each minor cycle contains thirty-six 3.77-mc pulses. The other part of the clock produces the pulse needed in other parts of the machine. Since the meaning of the pulses in the acoustic lines in the internal memory depends on their position relative to the start of the major cycles, the loss of the first part of the clock is equivalent to the loss of the contents of the internal memory. The two parts of the clock are independent and only a power failure or trouble in part one causes the loss of part one of the clock.

Central Control

Orders to be executed are transferred in sequence to central control and are executed a half order at a time. A complex set of selection rules for the operands are used to minimize the total operation time. The transfer of half orders into and within the central control is transfer weight counted. When an operand address has been used a check is made against the original address as received in central control. The operation code in the half order is used to select the matrix for the operation and this in turn produces a "tag" which is compared with the original operation code. Checks are provided so that if the machine fails to operate in the proper sequence of half orders, the machine stops. Two other operation checks are provided in the central control. A substitution is an operation modifying a half order which is to be used immediately and for its effective use the third operand must be selected before the fourth. Failure to do this stops the machine as well as certain mis-operations on branch operations.

The Arithmetical Unit

The objectives of the arithmetical unit should be clear from the description given above of the list of orders used in the machine. There are three main registers A, B, and C in the arithmetical unit, having both static (parallel) and serial methods of storing and handling information. Arithmetical operations are performed in parallel, multiplication is by repeated addition. The registers in the arithmetical unit have storage properties and their addresses may be used as addresses of operands in orders. In addition to transfer weight counts on the transmission of words, the arithmetical unit contains facilities for checking all the operations by means of checking formulas involving the sum of the digits of an operand when expressed in radix 32. These facilities involve an arithmetical check count register, a primary transfer register, a secondary transfer register and a serial multiplier which operates on the above mentioned checking quantities.

The Internal Memory

This consists of 36 acoustic delay lines each containing 32 words. An address in the memory consists of two parts, one of which identifies the line, the other the time during the major cycle at which the word in the address becomes available.

The Console and Directly Connected Printer

Words and orders may be entered into the machine from the console from banks of keys provided for this purpose. Transfer weight counts are formed in this process. The console contains lights such that if the machine stops, the operation point of the machine is clearly indicated and in the case of certain alarms lights will indicate the cause of the trouble. Results may also be printed out during a computation on the directly connected printer. There is a "vane check" on this printer also as well as the usual transfer weight count check.

Reliability of a Large REAC Installation

BERNARD LOVEMAN†

I. Introduction

PROJECT CYCLONE at the Reeves Instrument Corporation is under the cognizance of the Bureau of Aeronautics of the Department of the Navy. The primary function of Project Cyclone is the development and operation of a Guided Missile Simulator and the establishment and operation of a Simulation Laboratory. Problems in aeroelasticity, engine control, aircraft stability, dynamics, and navigation have also been studied with the aid of the computing facilities of the Simulation Laboratory.

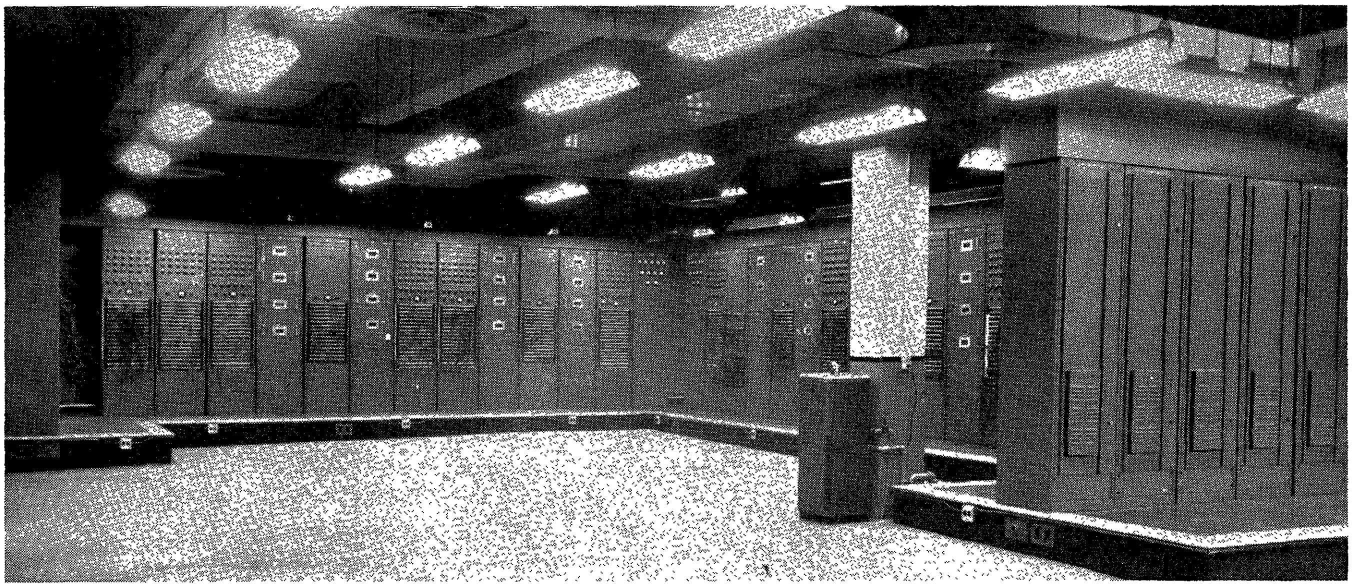


Fig. 1—Project Cyclone Simulation Laboratory.

Early in October, 1952, a new large simulation laboratory was put into operation and subsequently subjected to exhaustive acceptance tests. Fig. 1 shows an over-all picture of the computing equipment of this laboratory.

The power supplies and voltage regulators are located in a room behind the computer laboratory. The essential elements of this analog installation are the Reeves Electronic Analog Computer (REAC) and the Reeves Computing Servomechanism. A description of the individual units, their interconnection, and their use in solving problems was presented at the Joint Western Computer Conference.[1]

In June, 1953, Project Cyclone acquired a medium-sized, medium-speed digital computer with magnetic drum storage, the Elecom 100. This computer is being used primarily to obtain checks on the analog computer solutions.

The problem of maintenance and reliability becomes more and more serious as the complexity and size of a computing facility increases. The size of an analog computing installation may be evaluated in a variety of ways. Criteria useful in this evaluation are: (a) the number of computing amplifiers; (b) the number of vacuum tubes and crystals; and (c) the power consumed.

TABLE I

Breakdown of Components in Simulation Laboratory

Tubes: Type	Tubes: Number	Components	Number
6SJ7	1,222	Crystals (IN38)	392
6L6	536	Neon bulbs (NE51)	392
6SL7	451	Synchronous converters (vibrators)	483
5691	91	Scale factor potentiometers (0.1 per cent linearity)	446
6SN7	78	Servo potentiometers (0.025 per cent linearity)	222
6AS7	66		
6B4	60	Precision wire wound resistors (.05 per cent): 1 Megohm	1,423
6SK7	52	500 Kilohms	15
6AL5	52	250 Kilohms	430
12AT7	44	100 Kilohms	603
12AX7	39	Polystyrene computing condensers	115
6AU6	23	Film resistors (1 per cent)	1,948
5R4	18	Wire wound resistors (1 per cent)	734
6H6	12	Carbon resistors (5 per cent)	11,865
Misc	8	Condensers	5,270
		Relays	544
Total	2,952		

The new Simulation Laboratory contains 404 dc amplifiers, 2,950 vacuum tubes, 392 crystals, and consumes about 35 kw. A detailed breakdown of the components in the laboratory is given in Table I.

† Project Cyclone, Reeves Instrument Corp., New York, N. Y.

[1] L. Bauer, "New laboratory for three-dimensional guided missile simulation," *Proc. Western Computer Conference*, pp. 187–195; 1953.

From users' point of view, the most significant criterion of size is expressed in terms of problems that can be solved. At Project Cyclone problems involving equivalent of as many as 91 first-order ordinary differential equations, linear or nonlinear, can be solved.

At this point a comparison of analog and digital computers is in order. The reliability demanded of parts in digital computers will be appreciably greater than that of corresponding components in their analog cousins. This can be shown by tracing the steps involved in obtaining the solution to a complex problem. For analog computers the first step is to decide on the simulation technique and code the problem accordingly. A similar procedure is followed for digital machines where a numerical method is selected and the problem coded. The second step is to insert the code into the computer. On analog computers, this is accomplished by interconnecting the appropriate components with patch cords; on digital computers, an input device is provided. The third step for both analog and digital computers is trouble-shooting the setup or code checking.

Here the similarity ends; for in a dc analog computer the time for one solution is usually 60 to 90 seconds. On the other hand, for most digital computers, especially where a large number of variables must be printed out, a single solution may take from one to several hundred hours. A component failure results in a significant difference because of this time factor. In the analog computer, a failure invalidates only one minute of computing time. In the digital computer, only error-free, permanently stored results can be salvaged.

A second important difference arises from the manner of detecting intermittent failures. It is well known that because such failures are more difficult to locate they are far more serious than sudden total failures. In the REAC most voltages are "smooth" functions of time which are plotted on recorders or plotting boards. The presence of intermittent failures will be evidenced by irregularities in these functions and thus can be observed easily. In addition, the true overload system (see section V) detects all sudden surges or pulses. On the other hand, the computing voltages in a digital computer are pulses and the machine output consists of tabulated values. An intermittent failure may be interpreted by the computer as a signal and cause a circuit to function improperly. All subsequent calculations are invalidated. The overflow circuits may detect this failure. However, as a general rule analysis is required in order to insure that results are free of error caused by such spurious signals.

Two additional differences in operating procedure should be mentioned. The profusion of results from an analog computer facility requires that one or more operators be in attendance to monitor and process the solutions. For example, in one large problem the solution rate reached 200 runs per day. On the other hand, a digital computer may run unattended for many hours. Finally, an important feature of some analog installations including Project Cyclone is the flexibility of interconnection. This enables the operator to select the most suitable simulation setup and also the optimum number of components. From the standpoint of reliability, flexibility assures that no extraneous components will be used in the problem. This is in contrast to the digital computer or the prewired analog computer where all elements of the system must be operating to obtain solutions. Furthermore, the Project Cyclone Simulator may be divided into several parts in order to solve smaller problems. As many as four different problems have been on this computer at one time.

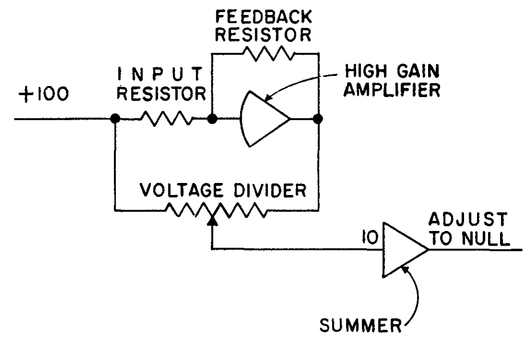


Fig. 2—Amplifier gain calibration circuit.

II. Acceptance Test

The Acceptance Test for the new Project Cyclone installation was divided into two phases. The first consisted of equipment checks. Every computing element in the system was performance tested and checked for accuracy. A few interesting tests follow.

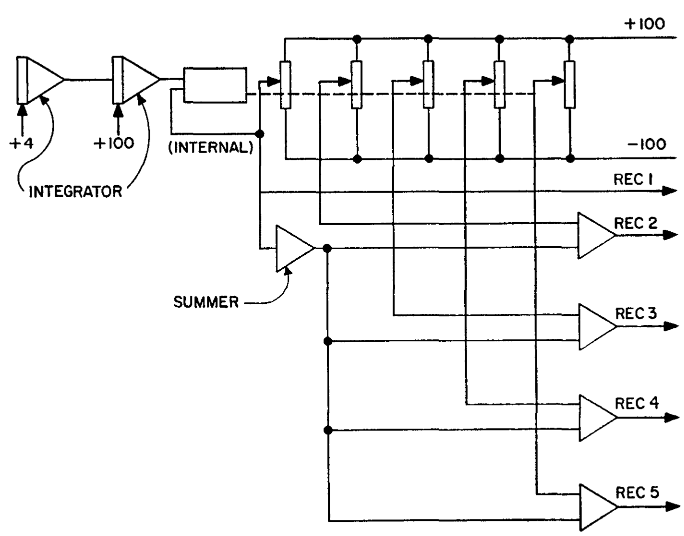


Fig. 3—Circuit used for checking servo-multiplying potentiometers with followup potentiometer as standard.

Fig. 2 is a schematic of the bridge circuit used to measure the input gains of every computing amplifier. A Leeds & Northrup 0.001 per cent voltage divider was used as the standard. The maximum permissible error in gain was 0.05 per cent. 94 per cent of the gains were within 0.025 per cent of their nominal value; 70 per cent were within 0.012 per cent of their nominal value. The mean of the absolute values of the errors in gain was 0.01 per cent.

The followup potentiometer of each servo was checked with a similar bridge. The setup used to check the multiplying potentiometers with the followup potentiometer taken as a standard is shown schematically in Fig. 3. The error voltages were plotted on a Brush recorder for the complete excursion of the potentiometer wiper. Maximum acceptable deviation was 0.05 per cent.

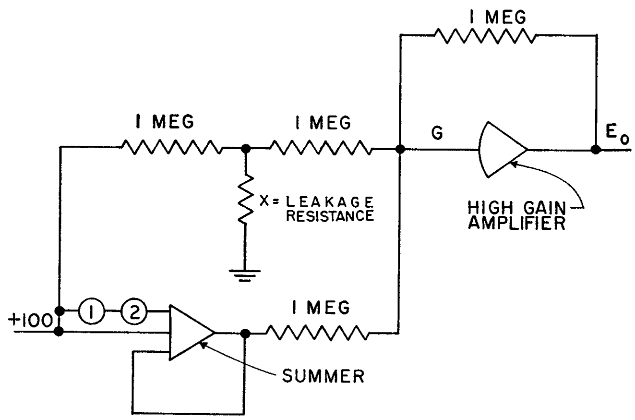


Fig. 4—Circuit for measuring interconsole leakage resistance.

The circuit used for measuring the leakage resistance to ground of interconsole leads is shown in Fig. 4. Potentiometers 1 and 2 are used to null the setup before the interconsole resistance is plugged in. The leakage resistance is calculated from the formula,

$$X \text{ (MEG)} \approx \frac{25}{E_0} \cdot$$

The equation is valid for $X > 100$ meg or $E_0 < 0.25$ v. The minimum acceptable leakage resistance was 1,000 megohms.

The second phase of the acceptance test consisted of a computing check. A highly complex nonlinear simulation problem, for which a numerical solution was available, was solved. An indication of the complexity of this problem is given by the fact that a single check solution took 75 hours of computing time on the IBM CPC. The problem required 304 computing amplifiers, 30 multiplying servos, 11 resolvers, 14 diode-function generators, and associated equipment.

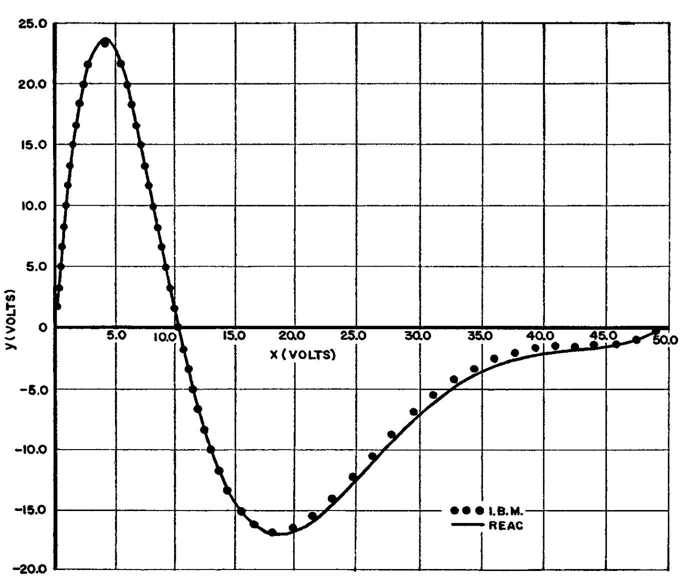


Fig. 5—Results of acceptance test.

The REAC plots for 26 variables were checked against corresponding numerical solutions with excellent results. Figs. 5 and 6 show a comparison between the REAC and numerical solutions for two variables.

III. Reliability

Prefatory Remarks

In an analog computer two classes of maintenance service are required, namely, adjustments and repairs.

Adjustments are provided in order to optimize the performance of the equipment as the components change characteristics. Some examples of adjustments are zeroing computing amplifiers, setting gain and damping controls of computing servomechanisms, and balancing limiting amplifiers. These adjustments are made as required, once a week or less frequently.

Repairs are required whenever a failure occurs. Failures arise from three causes:

(a) defective parts; (b) operator abuse; and (c) component deterioration.

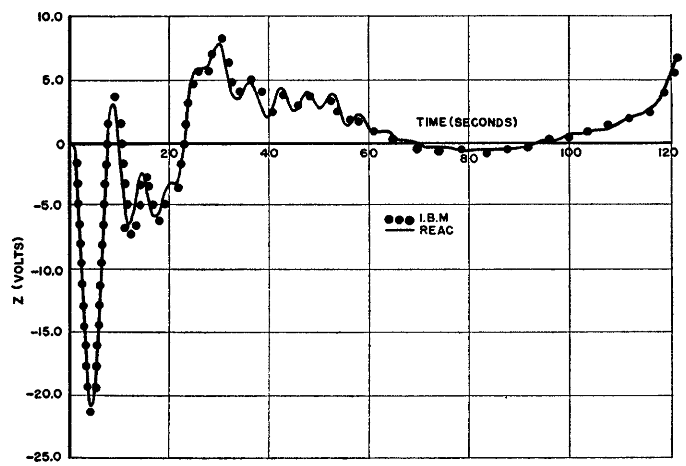


Fig. 6—Results of acceptance test.

In general, the number of failures may be reduced by adequate inspection and testing to eliminate defective components wherever possible before installation in the system. The procedure at Project Cyclone is to inspect and measure all components when they are received and then check them operationally after installation. Occasionally manufacturing defects appear after appreciable operating experience. One example is a potentiometer card that contained a scribe indexing mark under the winding which produced a weak spot. After 1,500 hours of operation, the winding broke. A change of manufacturing procedure has eliminated this type of failure.

Skilled operators are required in order to minimize failures and downtime. Operators may cause failures by carelessly overloading components. The following examples illustrate this point. If the reference computing voltage (plus or minus 100 volts in the REAC) is applied to the arm of a potentiometer which is at a low setting, excessive current will flow through the turns and the winding will open. The operator may also damage the mechanical parts of a computing servomechanism. The servo motor is coupled to the potentiometers through a dry-disc clutch with mechanical stops provided to prevent damage to them. If the voltage input to the servo is excessive, it will hit the stops; and if this voltage is maintained, the clutch will slip and be impaired. Finally, the operator must be able to distinguish between malfunction of the equipment and problem errors.

Failures due to deterioration will be reduced by conservative design (discussed in section V) and operating policy. Operating procedures which contribute to extending the life of computer components are time delay between the application of filament and plate voltage and adequate cooling.

Definitions

Two complementary concepts are used to evaluate the reliability of this analog installation. The first is based on the service-free hours and the second on the number of failures. These ideas are embodied in two coefficients defined below.

The first of these is the *reliability coefficient,*

$$R = 100\left(1 - \frac{S}{T}\right)$$

in per cent, where S is the service time and T is the total scheduled operating time.

Service time is the repair time spent on failures arising from both component deterioration and operator abuse. It includes, in addition to equipment downtime, repairs performed in nonoperating time. Under nonoperating time are subsumed checkout periods, Saturday, and time spent on chassis replaced by spares. Service time does not include operator time spent in locating trouble; nor is the time spent in making adjustments included.

Because this is a flexible simulator, the failure of a component in many cases signifies that only a part of the system is inoperative; other problems will continue uninterrupted. This is not reflected in the factor, T.

The second coefficient is the *failure rate,* defined as the percentage of failures of a specified type per hundred operating hours.

The coefficients are evaluated monthly.

IV. Reliability Data

Project Cyclone's weekly operating schedule has been set at fifty hours by the Bureau of Aeronautics, Department of the Navy. This is achieved in a ten-hour day, five-day week. During its first year, the installation has been in operation 2,527 hours. The data presented in this report have been accumulated during this period.

When the new Simulation Laboratory was installed, it was decided that, with the exception of operator checks, only breakdown maintenance would be provided as long as problems were plugged in; upon completion and unpatching of a problem, time would be allocated to check components and repair all defective parts. This course of action was adopted because the extent of a preventive maintenance program for a large analog simulator had not been established.

During the first twelve months, there have been six partial system checkouts. The time between checkouts of any cabinet varied from two months to a year. Operator checking is performed daily. It consists of examining all amplifier outputs with the computer in Balance Check, i.e., with the inputs to the amplifiers disconnected. Typical of amplifier failure is a noisy output or inability to zero the amplifier. Defective amplifiers are repaired immediately or replaced with spares.

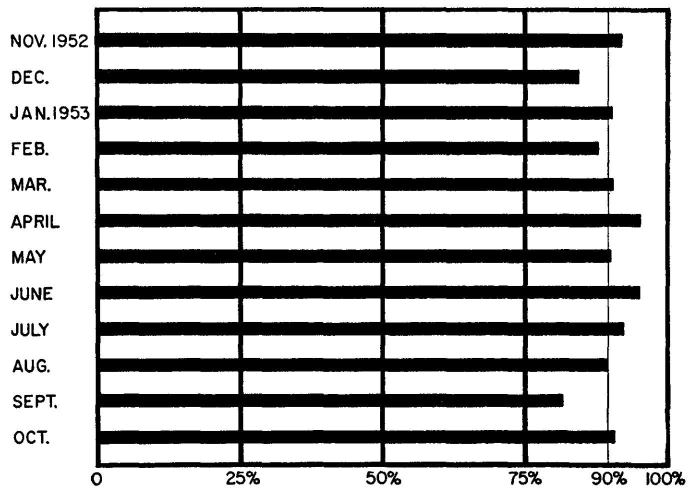


Fig. 7—Monthly reliability coefficients for Project Cyclone Simulation Laboratory (average for first year 90.6 per cent).

The monthly reliability coefficients are plotted in Fig. 7. The average reliability coefficient for the first year exceeded 90 per cent. The total service time was 234 hours. Approximately 75 per cent of this was expended in repairing failures caused by component deterioration; the remainder was due to defective parts or operator abuse.

TABLE II

Selected Failure Rates for the First 12 Months (2,527 Hours) of Operation

Component	Number	Per Cent Failure per 100 Hours
6SJ7	1,222	.38
6L6	536	.19
6SL7	451	.29
12AU7	200	.30
5691	91	.22
Vacuum tubes of all types	2,952	.38
Vibrators	483	.057
Deposited film resistors	1,948	.026
Servo motors	40	1.08

The failure rates for certain selected components are given in Table II. The criterion for this selection was a sufficient number of failures from which a significant estimate of the failure rate could be made. It is interesting to note that vacuum tubes constituted 63 per cent of all failures.

The components listed in Table II are largely in the

404 computing amplifiers, namely, 69.6 per cent of the vacuum tubes, 81.3 per cent of the vibrators, and 89.7 per cent of the deposited film resistors. Therefore, a careful scrutiny of the failures in these amplifiers is desirable. Three hundred and nine amplifiers (76.6 per cent) have not had a single failure during the first twelve months of operation of the Simulator Laboratory. Conversely, ninety-five have had one or more failures. These ninety-five amplifiers had a total of 207 failures of which 164 were vacuum tube failures. The average repair time for a failure was slightly less than twenty minutes.

At least one tube failed in eighty-five amplifiers. The distribution of vacuum-tube failures in these amplifiers was as follows: (a) forty-five had one tube failure; (b) twenty-three had two; (c) nine had three; and (d) eight had four to seven.

Not one vibrator in a computing amplifier failed during the 2,527 hours covered by this report. However, seven vibrators located elsewhere failed in this period; five of these were in meter amplifiers and two in servo amplifiers. Five of the vibrator failures occurred within the first 300 hours of operation.

The deposited film resistors which failed have been traced to a defective lot. The quarterly failure rates for deposited film resistors were .15 per cent, .046 per cent, .033 per cent, and .033 per cent. The defective resistors are reflected in the higher failure rate earlier in the year.

The first year of operation revealed one design weakness. Certain two-phase, 60-cycle servo motors are operated at twice rated voltage in order to drive a greater load with improved dynamic performance. The resulting stress causes a gradual deterioration of the insulation between windings which finally leads to a short circuit between stator field coils. The cure for this is either improved insulation or a larger motor operating at rated voltage.

V. Design Features

Present Features Tending to Increase Reliability

The effect of component deterioration is mitigated if a wide variation in characteristics can be tolerated. In a dc analog computer this is accomplished by using large amounts of feedback. The dc amplifier, for example, has 90 db feedback with a gain of one.

A second feature that contributes to reliable operation is the introduction of a true overload-indicator system, rather than one that indicates overloads at some arbitrary value. Both visual and audible indication of overload is provided. This enables the operator to take remedial action immediately to relieve the arcing across the vibrator contacts and thus extend the life of the chopper. This conclusion is substantiated by the fact that no vibrators in computing amplifiers have failed.

The tubes in a dc analog computer operate in Class A as in normal communication applications. The life expectancy should equal or exceed that achieved in these applications, because the tube operating conditions are conservative, e.g., the quiescent plate dissipation of the 6L6 output stage of the computing amplifier is only 3 w.

The use of plug-in interchangeable computing amplifiers reduces both service and downtime. The number of spare amplifiers necessary to maintain this installation is less than 3 per cent of the total.

Design Program

The first year of operation has demonstrated that the simulator has a very high reliability. Therefore, only a limited preventive maintenance program is required. The cost of this program will be carefully evaluated relative to the improvement in system performance.

The present system of patching precludes the possibility of checking permanently wired components when a problem is wired in. This difficulty can be overcome by installing a suitable prepatch system. At the time (1951) the present Project Cyclone Laboratory was designed, no large prepatch facilities were in operation. Since experience with this idea was very limited, the proven telephone jackbay was selected. With a prepatch setup the problem can be removed readily if a failure is suspected, and a test circuit substituted for it. If the necessary repairs take too much time, the problem board can be plugged into a spare computer. Thus, prepatch adds to the flexibility of the installation and downtime is reduced.

Acknowledgment

The author wishes to acknowledge the assistance rendered by his associates in writing this paper.

Discussion

M. L. Aitel (RCA): Could you comment on the life and stability of the crystal diodes and of the NE51 neon lamps? How were the neons used?

Mr. Loveman: Both the crystal diodes and the neons are used in the true overload circuit that I discussed. Normally, that is when the amplifier is not overloaded, the crystal does not draw appreciable current, and the neon is out. During the first year (2,527 hours), eight crystal diodes and five neon bulbs failed.

A. S. Marthens (Department of the Navy, Bureau of Ships): Were amplifiers tested individually before installation in the computer? If they were, what sort of test was made, and how many failed the test?

Mr. Loveman: All the amplifiers were individually inspected and tested as they were produced. All components and solder joints were inspected. Each amplifier was buzzed out before power was applied to it. Each one was then checked for balance, for noise, and for proper operation as a summer and as an integrator. After installation the amplifiers were tested operationally.

The foregoing tests were all performed before the equipment was turned over to Project Cyclone. No information on amplifier failure during this period is available.

S. T. Smith (Hughes Aircraft): What was the nature of tube failures? Loss of emission? Failure of cathode heaters? Tubes gassy? Defective assembly?

Mr. Loveman: Until very recently we did not analyze the reasons for tube failures. We are doing that now but as yet do not have the desired information.

Computer Performance Tests Employed by the National Bureau of Standards

S. N. ALEXANDER AND R. D. ELBOURN†

Introduction

DURING THE PAST five years the National Bureau of Standards has collaborated with five government agencies in technical programs that involved the performance testing of complete electronic digital-computing systems. Devising a meaningful evaluation procedure proved to be a surprisingly elusive task. Because digital computers are highly integrated assemblies, there is not much value that the prospective user gains from a detailed check on each of the individual facilities and units.

It appeared far better to concentrate on over-all system testing, since the tests could readily be programmed to use all the facilities and units rather thoroughly. Moreover, if the situation justifies the sophistication, the program can be arranged to permute the number patterns, storage locations, and the computer operations, and thereby scan over a variety of combinations in search of marginal ones.

The performance tests devised jointly by NBS and the Bureau of the Census for the acceptance testing of the first UNIVAC system were planned along these lines but did not include all the refinements. This initial experience was described briefly at the Joint Computer Conference in 1951. Since testing the Census UNIVAC, NBS has employed nearly the same tests in accepting four more UNIVAC systems for other government agencies. The purpose of this paper is not only to bring the record up to date by describing how these machines performed in their tests, but also to examine what we have learned about the general problem of acceptance-testing machines such as these.

Design of the Tests

When we undertook to design acceptance tests for the first UNIVAC system we had very little to guide us in the way of experience with the kind of performance which could be expected of such machines; however, the Bureau of the Census did have some estimates of the minimum performance which would be necessary for the machine to be economically competitive with the conventional tabulating-card methods of doing Census work. The design of the tests leaned heavily on these estimates of minimum adequate requirements and on observations of the performance that was being achieved during the manufacturer's debugging tests. We hoped to learn from subsequent operating experience how the tests should be modified for future machines.

† National Bureau of Standards, Washington, D. C.

Census officials felt justified in accepting a machine that would be in productive operation at least 50 per cent of the time, so downtime in the test was limited not to exceed running time. But for reasons which will be discussed later, the downtime observed in a brief test is usually a grossly optimistic prediction of the downtime to be experienced in use.

The more significant rules of the test were derived from the knowledge that the principal task of the Census UNIVAC for a year or more would be a tabulating job in which the time required to tabulate the data from one reel of magnetic tape would be 20 minutes. This meant that Census wanted a machine that would usually go through a 20-minute problem satisfactorily. Four successes out of five tries was selected as a nominal figure which would not result in excessive rerunning. To decide how many repetitions of a 20-minute test routine should be required in order to demonstrate this probability of successful completion, Census statisticians proposed following an adaptation of standard sequential sampling plans based upon the proportion of defectives. According to the theory of this plan a machine would pass the test 99 times out of 100 if its probability of running through the 20-minute routine successfully were 0.9, but it would fail 99 times out of 100 if this probability of success were only 0.7.

The theory, however, assumes that the basic probability of running the 20-minute routine successfully is independent of time or of machine performance in other 20-minute periods. Actually supply voltages and component values are continually drifting and changing, so this probability is not constant. It is not clear how seriously this impairs the probabilistic interpretation of the test. Each 20-minute test unit is charged with at most one defect. This avoids the problem of how to score the multiple errors that might occur at one breakdown and gives a greater measure of independence to the defects.

The procedure of a sequential test is very simple. It can be governed by a chart, like Fig. 1, of the number of defective test units versus the number of test units completed. As long as the graph stays between the parallel sloping lines, testing is continued. If it crosses the lower line the machine is accepted, or if it crosses the upper line the machine is rejected. The advantage of a sequential test is that it reaches a decision relatively quickly if the quality of the machine is very much above or below specification. Testing tends to become as protracted as in a fixed duration test of comparable discrimination only if the quality of the machine is actually marginal. If no errors occur, this particular test will be passed in

19 test units. Each defective test unit must be compensated by 5 or 6 additional nondefective test units if the machine is to be accepted. If on the average one unit in five were defective, the expected duration of the test would be 76 test units.

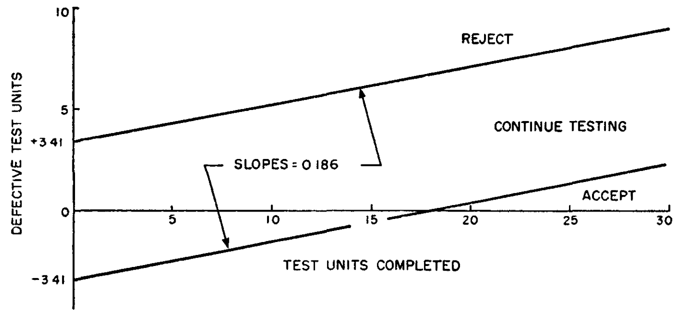


Fig. 1—Chart for controlling the sequential test. Slopes and intercepts are determined by the specified probabilities.

Two objections to this testing procedure immediately arose. First, in a machine like UNIVAC with extensive built-in checking circuits and diagnostic facilities in the Supervisory Control not all errors are equally serious. In many cases all the operator has to do is to clear certain registers and cause the last instruction to be repeated in order to resume the problem successfully. Second, errors of this less serious kind actually occurred too frequently in reading magnetic tape for the machine to be expected to perform four error-free test units out of five. Therefore, two categories of defective test units were recognized, "major defects" and "minor defects." Major defects were controlled by the sequential plan as already described and minor defects were limited to no more than one-third of the test units completed. It seemed impossible to list all the kinds of malfunctions which could occur and to classify them as major or minor defects, so for definiteness the following somewhat arbitrary rule was adopted. A test unit was charged with a minor defect if the built-in checking circuits stopped the computer and the operator was able to cause the unit to be processed to completion correctly by using only the external manual controls of the machine. Even restarting the routine was permitted. A test unit was charged with a major defect if any internal adjustment or even an internal inspection was performed, because this normally would require the services of a maintenance engineer or technician. Any error undetected by the built-in checking but caught by the programmed checks would also have been a major defect but this never happened.

The central computer and its associated magnetic tape input-output units were given two tests. Each was conducted according to the procedure just described, but different test routines were used. The "general test" checked the internal computer operations exhaustively and made only minor use of the magnetic tapes, whereas the "tape reading and writing test" used the central computer only for control and for checking the tape operations.

The 20-minute test unit of the general test was made by performing a 10-minute routine twice. A subroutine of about 2,590 operations, which used and checked every available internal operation was repeated 404 times in $8\frac{1}{2}$ minutes. The remaining $1\frac{1}{2}$ minutes were occupied by solving a heat-distribution equation, doing a few input-output operations, and sorting 10 numbers into numerical order.

There were two versions of the tape reading-and-writing test. The first, used on the Census UNIVAC, required 2,000 blocks or about 1.4 million digits of alphanumeric data to be written on a tape, then read back and compared with the original data. After that it required writing and reading over one spot of tape for 700 passes to check for wear. There were also checks of various sequences of writing, rewinding, and reading and of the possibility that a tape might drift while supposed to be standing still. In all, the test unit lasted 25 minutes. Of this, 7 minutes were occupied by the reversals required for the 700 passes over one block.

When operating experience on the Census machine revealed that by far the most frequent errors were simply tape-reading errors presumably caused by dirt or by bad spots in tape, the tape reading-and-writing test was changed to require a maximum of straight reading and writing. In the new test 1,350 blocks of data were copied from one tape unit onto a second, then from the second onto a third and from that onto a fourth. The original and the fourth reels were then transferred to different units and compared. A total of 6,750 blocks were read, or over three times more than in the earlier test. When time occupied only by rewinding was declared time out, this routine lasted only $15\frac{1}{2}$ minutes. The constants in the sequential procedure were, therefore, changed to make the total running time required to pass the test as long as it would have been if the test unit had been 20 minutes.

There were also tests for each of the separate auxiliary machines of the system, but the contents of these was described at the first Joint Computer Conference, and they involved no distinctive philosophy of testing so they will not be discussed here.

TABLE I

General Test

Machine Number	1	2	3	4	6	Total
Date passed	Mar. '51	Nov. '51	Apr. '52	Aug. '52	May '53	
Major defects	0	0	0	0	0	0
Test units run	19	19	3	3	18	62
Minor defects	2	3	1	1	3	10
See note*	1	0	0	0	1	2
Running time, min.	380	381	60	60	363	1244
Downtime, min.	17	3	†	†	4	24

* Minor defects that involved central computer malfunctions.
† Less than one minute.

Table I shows the performance of the five machines on the general test. There were no major defects at all. Minor defects were only about half as frequent as the allowed one test unit in three, and only two of these were caused by errors in the central computer. Most of the minor defects were tape-reading errors in spite of the little use made of the tapes on this test. The downtime was ridiculously low. After the first two machines had been tested, it was obvious that the performance of the central computer far exceeded the test requirements. Therefore, on the third and fourth machines the general test was curtailed to three test units on the condition that these be run immediately after the tape test was passed. Only one minor defect was permitted. The buyer of the sixth machine did not agree to this curtailment so the complete test was run on it.

Table II shows the performance of the five machines on the tape reading-and-writing test. There were two major defects caused by defective tubes and two other major defects which were quite trivial. They were merely failures of a fuse which inadvertently had been chosen too small for a certain peak load condition.

TABLE II

Tape Reading and Writing Test

Machine Number	1	2	3	4	6	Total
Date passed	Mar. '51	Feb. '52	Apr. '52	Aug. '52	May '53	
Major defects	0	1	0	1	2†	4
Test units run	19	32	25	32	39	147
Minor defects	6	8	8	9	4	35
See note*	4	8	7	7	3	29
Running time, min.	465	477	382	498	609	2431
Downtime, min.	103	126	116	124	37	506

* Minor defects that involved input-output malfunctions.
† Not significant—improperly rated fuses.

Minor defects, however, present another story. They approached the one-third limit on all but the last machine and were predominantly tape-reading errors. The downtime reflects both confusion and rerunning caused by the reading errors. The performance figures on the first machine are not really comparable with the others because it passed a less severe test. If significant improvements had not been introduced into the second and later machines, it is doubtful whether they could have passed the stiffened test. The sixth machine with only three tape errors in 39 test units appears, finally, to reflect the cumulative effect of many engineering improvements in the tape system.

Critique of the Test Design

Within the first few minutes of running on a test such as this, all the basic operations of the computer have been used and checked; moreover, the running time of the routine verifies that the speed of the computer is up to specification. All the rest of the test is an attempt to measure the machine's reliability. The quantity which a purchaser would like most to learn from the test is the average proportion of good running time that he can expect in use, but a test measured in hours or tens of hours is inherently unable to measure this quantity. It is far too short to collect a representative sample of the serious breakdowns that affect the long-term efficiency. Indeed, if one of these occurs during the test, the manufacturer must be permitted to discontinue the test, repair the machine, and resubmit it for testing. Thus an acceptance test is biased in the direction of containing no breakdowns or only minor ones. Moreover, during the test the machine is under the care of the team of specialists who built it, and all preventive maintenance is done before the test.

A downtime limitation is retained in an acceptance test primarily to keep it from turning into a protracted debugging session. From the fluctuations in the good running time that appear in the logs of operating computers it appears that several months are needed to get an accurate measure of the average proportion of good running time. If a purchaser really requires a measurement of long-term efficiency before final acceptance, then he would be better served by some kind of rental-purchase plan with a performance guarantee rather than an acceptance test.

Although tests such as these cannot guarantee long-term operating efficiency, they establish something of distinct value. They prove that the machine overrides very well the many disturbances that are always present, crosstalk and ambient electrical noise, line voltage fluctuations, dust on the magnetic tapes, and the like. To achieve this in so complex a device requires a lot of painstaking engineering debugging to locate and correct marginal tolerances in the system. On the first machines of a new design this debugging may require many months. Until the marginal tolerances are fairly well eliminated, long error-free runs will hardly occur. The requirement of substantial error-free running periods in an acceptance test is, therefore, a powerful incentive to thorough debugging. In fact, the real purpose of these tests is to insure that, if any intermittent or sporadic errors remain in the system, the intervals between them are at least long enough to allow useful work to be done.

Repeating the same 2,500-operation routine 808 times per test unit might well be criticized for not utilizing fully the opportunity to test various combinations of numbers, operations, and storage locations; however, the total number of possible combinations is so vast that even a perpetually changing routine can test only a minute fraction of them. The only way in which one is likely to find any particularly troublesome combinations is through a much more intimate knowledge than we had of the details and error predispositions of specific circuits in the machines. For this reason the manufacturer's own debugging routine called NOF was a more searching test of these particular machines, but we preferred to use at least a different routine than had been used for most of the debugging. The value of these tests as acceptance tests was certainly not impaired by having some magnetic tape operations in the general test, which was primarily concerned with computational ability, but a clearer picture of the central computer's performance would be presented if they were omitted.

Although the choice of a 20-minute test interval was originally based upon a particular job of the Bureau of the Census, it seems that for safety one would usually break longer problems into units no longer than this by writing on a tape all the information needed to restart the problem there. No other purchaser presented reasons for using any significantly different duration of test unit.

The allowance of about one major defect per five test units was somewhat arbitrary, and it was indeed gratifying to discover that the actual frequency is perhaps one-tenth of this allowance. If a purchaser requires assurance that major defects in his machine are really this infrequent, the test would have to run for hundreds of test units and thus would be impractically long.

Because minor defects were the only factor in which performance did not greatly exceed specifications, it might have been more logical had the duration of the tests been controlled by a sequential plan on the minor defects. An advantage of this scheme is that, if the testing becomes protracted because the performance is marginal, the manufacturer has an incentive to terminate the test, rework the machine to a higher level of performance, and then resubmit it in the expectation of running off the test more expeditiously.

This raises one of the thorniest problems in the testing procedure, namely the permission to terminate a test and restart. It is not the purpose of these tests to reject a machine finally but rather to have it brought up to a useful level of performance and then accepted. For this reason the manufacturer was given the option of discontinuing a test at will and resubmitting the machine after taking significant corrective action. Because the significance of corrective action often cannot be proved, there was a tendency not to require it; thus an undesirable element of "pitch-'til-you-win" was permitted. In this connection, a purchaser should not insist upon specifications that are too close to all the existing state of the art can offer because then a "pitch-'til-you-win" game can scarcely be avoided.

In summary one might say that it is by no means insignificant to require a machine to perform 40 million internal operations correctly or to read correctly a billion binary digits from magnetic tapes. These tests did require that the most frequent of the sporadic errors be reduced to about one per hour. With a few changes the tests could be somewhat improved, but the basic limitations on what they can do lie in their duration. They cannot be life tests of components nor can they measure long-term reliability accurately without being orders of magnitude longer. These were tests of the state of debugging of the machines when they were accepted, and for their purpose they were quite effective.

Discussion

C. E. Bergman (Melpar Inc.): How many pulses per linear inch on Univac magnetic tape?

Mr. Elbourn: Approximately 100 pulses per inch in each of 7 channels.

George J. Christy (State of California—Dept. of Education): In testing tape reading units are any tapes which have been prepared and aged for a period of a year or more ever used? If so, what have the results been?

Mr. Elbourn: This was not possible in the test. We did use in the later machines tapes which had been used in census tabulating work. These have been erased, cleaned up, and re-recorded.

L. W. Armstrong (Bureau of the Census): Please comment on the fact that the supplier was familiar with the detailed content of the test. Would any useful purposes be served by concealing the details of the test but giving knowledge of the general nature of the test?

Mr. Elbourn: When we ran the first tests, the only one able to prepare a magnetic tape and debug the routine by trying it on the machine was the manufacturer, so he necessarily knew the details of the test. This question relates to the possibility that the machine may be sensitive to certain sequences of pulses or operations. One can never test more than a tiny fraction of all the possible sequences, but perhaps the manufacturer would be more diligent in debugging for various sequences if he could not pretest the machine with the routine of the official test.

Arthur S. Marthens (U. S. Navy—Bureau of Ships): How many component items—tubes, fuses, resistors, etc.—were replaced between the first offering by the manufacturer and the acceptance?

Mr. Elbourn: We do not know, but we have no evidence that the replacement of tubes, resistors, etc.—ran any higher than the normal replacement rate during use. The most significant changes were on one machine that was first submitted in November and finally passed its tape reading and writing tests in February. In that time a large number of design improvements were incorporated in the Uniservo magnetic tape reading mechanism. These improved its performance considerably.

Experience on the Air Force UNIVAC

ROBERT KOPP†

THE PURPOSE of this paper is to discuss those factors the Air Force finds to have the most impact on electronic-computer reliability. In order to accomplish this purpose, I shall tell the story, in a general way, of the Air Force's experience with the UNIVAC for the past 18 months. During this time it has been under the sole control of the DCS/Comptroller, Headquarters, United States Air Force. I shall correlate these experiences with the reliability aspect of electronic computation.

It might be well to start with an understanding of what we tentatively consider to be the meaning of the term reliability. As a representative of the engineering group, I find that reliability means essentially one thing, and that is: Is the computer available to do computation right now? This may seem to be a harsh definition to engineering personnel, but it must be realized that our electronic computer is merely a tool for certain groups of people to get a lot of work done fast. You need only to talk to people in the programming group, the mathematicians, the operators, or your boss, to find that this definition is the one they go by. Remember now, I'm talking about a computation installation where a huge volume of work is to be turned out, which in turn is the justification for the electronic computer in the first place. I am not referring to a manufacturer's plant where a computer is under test, or to a laboratory where a computer is a research tool or experimental device.

An electronic computer is a complex system of equipment, and as such is dependent upon all the many components that constitute an installation. The computer proper may be operating, and could do work if a required part were available to fix it. Or, perhaps, the main power lines are at fault because a power transformer was burned up. The fact still remains that the computer cannot function. These two simple but very realistic examples illustrate the inclusiveness of what we consider to be reliability. If you start making distinctions, an endless number of definitions become possible. Our definition is the one more and more customers will be going by as well.

It is worth noting that the use of this criterion is of some historical significance for the computer industry. Until operating computers were available which were more than research instruments, the definition would have been irrelevant. The definition expresses correctly enough the belief that the reliability of a computer can be compared to the reliability of a tractor, or a bomber, or to a desk calculator. With the aid of such a definition, I can best convey information at our disposal.

† Mathematical Computation Branch, U. S. Air Force, Washington, D. C.

Before relating the specific factors that have affected the reliability of our operation, I am going to tell something about the circumstances that had a direct bearing upon these factors.

Starting approximately June 1951, Air Force personnel were sent to the Eckert-Mauchly Division of Remington Rand in Philadelphia. No training program existed at this time. The computer itself was being checked out by the test engineers and, therefore, was not available for use by the trainees. A technical manual was well under preparation which covered the Central Computer. It was this manual and the constant questioning by the Air Force trainees of the test engineers that made up the training that was received. From time to time, as could be arranged, lectures were given by the Eckert-Mauchly test engineers covering various phases of the computer. I would like to pay my respects to those test engineers who suffered the continuous bombardment of questions by the Air Force trainees during the period June 1951 through February 1952. They were most helpful, most accommodating, and most all-knowing. During December and January 1952, the computer was occasionally made available during the hours of 12–8 a.m. for simulated trouble-shooting. This was the first and only real experience we had on the UNIVAC prior to its installation.

In early 1952 the machine passed its acceptance test, and in February it was shipped to the Pentagon for installation by the manufacturer's engineers. This was the first UNIVAC to undergo this transformation. From February to June 25, 1952, installation, checking and limited operation went ahead, and on that date the machine was formally transferred to the Air Force; the test engineers returned to Philadelphia, and we were left with the UNIVAC. It was quite a sensation.

During the training period, and especially at the start, there were indications of slight uneasiness on the part of the Air Force trainees at the prospect of servicing the computer. As the strangeness of the equipment began to wear away another form of uneasiness prevailed, not by the trainees now, but by persons associated in various ways with the installation. This latter uneasiness, however, turned out to be a blessing in disguise, because it constituted a challenge which made us all the more determined that the UNIVAC would be serviced only by the Air Force group without any help from "outside." This was a bold attitude: in a more practical light, it might be considered plain foolishness. As it turned out, however, the attitude was vindicated by results.

From the start it was decided to operate the UNIVAC on a 24-hour-per-day, 7-day-per-week arrangement. The personnel consisted of 11 people to perform all servicing on the Central Computer, the Unitypers and

Uniprinters, and the test equipment used for servicing. It has been no mean task. Of the group of 11, two persons were assigned, in a general way, to the Unitypers and Uniprinters, one for the mechanical features and the other for the electronic features. This was a situation that came about as a result of the natural bent of the individuals concerned, rather than any other reason. Two were assigned, in addition to their engineering duties which always came first, to a combination of re sponsibilities which covered such things as procurement of spare parts, personnel administration, dealings with programmers, and scheduling of preventive servicing. To visualize the communication problem, remember that with three shifts a day, seven days per week, contacts were limited to fifteen minute lap-over periods and written notes. At least two of the eleven were required to attend the UNIVAC at all times, performing computer operation, preventive servicing, or trouble shooting, as circumstances required. These two persons were familiar with all phases of UNIVAC servicing and operation rather than being specialists in any particular aspect of the computer such as a power-supply expert. This policy of shying away from specialization resulted, we believe, in greater reliability in utilizing our minimum number of personnel and, in addition, offered an opportunity for widening their interests and increasing their versatility. As we look back at this staffing arrangement, we realize we were undermanned.

A big problem in computer engineering personnel is to have a pool of trained people to slip into niches left by any who may leave. Based upon our experience, it takes approximately eight months to train a qualified person, and by that I mean an electrical engineer, a physicist, or a mathematician with a physics minor and a course in electronics. At the end of this time, he is almost ready and has started to have confidence in his own ability, so that he can take the responsibility of an operating shift. The work of servicing equipment such as the UNIVAC could be seriously hampered if too many engineers were to leave at one time, or if too much specialization was permitted in their training.

Another big problem concerns working hours. Anyone entering the computer field must immediately resign himself to the fact that he is going to have unusual working hours for a good percentage of the time. In many ways this can be compensated for by differential pay, true overtime, and a personnel policy that places a value on the services of employees.

An important Air Force accomplishment that has affected reliability was that of obtaining UNIVAC spare parts. While this may seem to be a routine task, it took a good portion of the time of two people for the better part of a year. The first step in securing adequate spares was that of obtaining a list of parts that would serve for immediate procurement purposes. A spare-parts list fairly adequate for this purpose was made available at the time of installation. It is possible that some of the burden of procuring spares could have been detoured at this time by obtaining a package of spares from the manufacturer. This was not done. Instead, we bought as many items as possible on the open market. As a result, we obtained valuable experience in procurement difficulties and established an excellent system of government procurement through Air Force Headquarters.

A serious problem that arises owing to the lack of an up-to-date and complete parts list of everything used in the computer is the continuing discovery of new parts that have never been ordered. This is a most serious factor in effecting computer reliability, for usually the component required is one that requires special ordering. In these cases, emergency procurement from the manufacturer is called for. At the start of our operations this problem was handled extremely well. Certain of the vendor's personnel were authorized to accept emergency calls from us at any time of the day or night, secure the part from stock and either bring it down personally or see it on its way by the most rapid means. This system worked extremely well and due credit must be given to the vendor and the individuals responsible. After one year of this system's successful operation, this service was withdrawn. It is true that the service had not been used a great deal, and the manufacturer naturally discounted its importance to us. After losing computer time in a few emergencies, we suggested that the service be reinstituted. This has been done. We have concluded that this kind of service is an important aspect of reliability since it is the one link that permits a computer installation to function, in spite of oversights on the part of both vendor and customer.

At this point, it might be well to discuss what spare parts have to be stocked by an installation using a UNIVAC. At the beginning, the selection was based on many intangible factors. One of them was: given such an amount of money, how much can I buy? This question certainly does not provide a very helpful spare-parts philosophy. When going into large scale computation with electronic computers, one might just as well budget for a generous supply of spare parts, rather than a minimum supply, since the difference in cost represents only a small fraction of the over-all cost of a large scale computer installation such as the UNIVAC. The compensation for this difference in cost is increased reliability.

Another important consideration in the achievement of greater reliability is the electronics work and storage area. The original Air Force plans merely provided for bare space which could serve as a work and storage area. Cabinets, shelves and benches have been added continuously, and the layout modified many times. Original ideas as to arrangement have been discarded and new ones developed continuously to take advantage of experience. Since no model existed to serve as a guide, we have no qualms about discarding existing arrangements in favor of more efficient servicing.

Arrangement of this work area has been dictated by factors which originally were not visualized. From the experience gained in our tube testing program, for example, we have concluded that a set-up for preheating a large number of tubes at one time is necessary. This has required bench area which we originally had not contemplated. We will continue to make changes as experience dictates.

As may be expected, the components giving rise to most of the computer troubles are tubes. To date the Air Force's UNIVAC installation has replaced approximately 2,100 tubes of all types to remedy failures. In addition, approximately 1,300 tubes have been discarded because they did not pass specifications and were rejected before even being used.

Several choices can be considered for meeting the procurement requirement for tubes. One of them is to buy pretested tubes from the vendor, if such are available. An alternative is to do your own selecting and testing. We have elected this last alternative.

For the UNIVAC it has been found that the tubes of Manufacturer A have been much better than Manufacturer B, whose tubes have become gassy, and Manufacturer C, whose tubes have lost most of their emitting characteristics after 2,000 hours of operation. Actually, it was known that Manufacturer A's tubes were desirable, but at the time of procurement only B and C's were readily available. Following the experience with the tubes of B and C, only A's tubes have been in use. Steps have also been taken to work out a procedure for conducting our own tube testing on a wholesale basis. It has been found important in tube checking to allow sufficient time for the tube to warm up to temperatures approaching those in the computer. Tubes that will not show shorts under testing conditions may do so if allowed to reach operating temperatures existing in the computer.

The most curious type of defect exhibited by discarded tubes has been a G2-K short. Not a dead short, but rather one that varies from approximately 10,000 ohms to 400,000 ohms. This situation, of course, gives rise to many intermittent-type troubles in the machine, because different circuits would react differently to various values of shorts. In most cases, even a 50K short should not make any difference. However, in many cases, the short condition evidently takes on values considerably lower than this; of the order of 10K, and it is this condition that causes the intermittent and, consequently the most frustrating trouble. It has been the practice, until recently, to remove all tubes that show a short condition of the order of 1 megohm. This might seem like a terrific waste of tubes until we see the reason for it. An experiment was conducted to determine the nature of the offending shorted tubes. The envelopes of many of the most common tubes used in the UNIVAC, the 25L6GT, were removed and the deposition of cathode material on the second grid has been observed. It has been determined that this deposition is caused by excessive heat within the UNIVAC. This condition, especially prevalent during the hot summer months, has caused cathode material to be torn loose and deposited on neighboring elements, in this case G2. Therefore, once a tube shows a short between G2-K, we assume that the condition will get worse and discard the tube. We believe, from a long-range viewpoint that this practice has served to increase reliability.

The UNIVAC is cooled by drawing in outside-temperature air wihout refrigeration and using it for cooling. An intake fan driven by a 20 hp motor pulls in air at a rate that was designed to be 30,000 cubic feet per minute. After being blown through the central computer and power supply, it is exhausted by a fan driven by a 15 hp motor. Despite these design figures, which are considered adequate to afford satisfactory operating temperatures, we started burning out selenium rectifiers at a rather alarming rate. Investigation led to the disturbing fact that the airflow was 18,000 rather than 30,000 cubic feet per minute. When this was discovered, some temporary measures were immediately taken, and the flow increased to approximately 25,000 cubic feet per minute. A definite program is underway to increase this flow to at least the original design figures, and it is expected that this modification will be installed in the very near future. In the meantime, the extra replacements caused by this difficulty decrease the reliability of our computer. This is all part of the computer game.

A rather serious consequence of the heat problem in the UNIVAC was its effect on personnel operating in the work area. The Pentagon Building, where the computer is installed, is air-conditioned all year round, and the summer temperature is in the neighborhood of 72 degrees F. The temperature in our work area hovered at 88 degrees during the summer, and 82 degrees during the winter. This is hardly conducive to comfort, and measures have been taken to remedy the condition. We hope that this problem will be recognized and avoided in all future electronic computer installations.

A unique feature of the UNIVAC is its input power. Many of you know that the UNIVAC was built in Philadelphia, a city of many old traditions including two-phase power. Perhaps this explains why the UNIVAC requires two-phase power. Since three-phase power is available in most installations, the only practical (and here I use the term very advisedly) solution to this problem is to convert available three-phase power to two-phase power via that enigma of the power men, a Scott transformer. To date, we have burned up two Scott transformers, and the third is presently operating successfully, we believe, only because we have installed a blower system over it. We have a new transformer on order which is 25 per cent larger in rated power than the one we are now using. We hope that it will cure our power failure problems.

The types of troubles that occur in the UNIVAC have never lent themselves to any sort of a pattern, other than to say that there are some types of troubles that recur. I want to attempt to outline these difficulties starting from the more prevalent to a less prevalent type. A very common trouble is blowing of fuses in the Direct Current lines. The UNIVAC is endowed with an alarm circuit which causes a grasshopper-type fuse to blow and drop all Direct Current supplies at the same time. Little data has been gathered to determine the exact nature of these fuse blowouts, other than when a dead short was discovered. However, in all instances where measurements were made, no case revealed a higher than rated current to be flowing through the fuse in question. This has led us to believe that the heat problem previously mentioned has been one of the causes of fuse blowouts. The blowing of a grasshopper fuse is usually nothing more than an annoyance, but, in rare cases, can be serious when the supply is one of the clamping voltages, because then it may take with it many of the clamping diodes. This can be a rather annoying procedure, because it means checking every chassis that uses this voltage, and removing the bad diodes. This can easily consume the better portion of an 8-hour shift. At the beginning of our operation, we were consuming diodes at the average rate of $1\frac{1}{2}$ a day for all reasons, but mainly because of blown fuses. At the end of six months this usage had dropped to almost nothing, and was coincident with our increased cooling effort. Our present diode replacement consists almost entirely of removing doubtful diodes of manufacturer X, and replacing them by manufacturer Y, or the more positive, but less occasional, situation of removing a burned-out diode which can occur when a clamping voltage fuse blows. To date, we have consumed a total of approximately 230 diodes.

Components, such as resistors and capacitors, have had negligible replacement factors; also fitting in this category are most of the other items including relays, lever switches, transformers, and chokes. While it may appear from these statements that only a small stock of spare parts need be kept on hand, the question then comes up: Which ones shall we stock? Can we afford to take a chance on the other items, remembering that regular procurement may take as long as six months, and securing it from the computer's manufacturer not less than 8–24 hours depending upon the circumstances? It must be remembered that one never knows what components might cause trouble and be weak links in the computer installation. So, it appears that until computers are much more common than they are today, it is almost mandatory that the computer user stock his own spare parts in a quantity allowing for self-sufficient operation.

We have found three kinds of problems associated with the use of magnetic tapes. The first problem is related to their shipment. The density and composition of a reel of tape is such that movements during shipment can cause layers to slide with respect to one another and produce folds in the tape. Each fold is a source of read trouble when the tape is used. This problem is being reduced by strict packaging methods. The second problem is an accumulation of dust and dirt on the tape surface. This is a cumulative process, and it is difficult to say when it starts becoming serious. However, there is no question that dirty tapes cause read-and-write errors that are time-consuming. The third type of problem is tape breakage, usually caused by some mechanical difficulty. Tape breakage requires tape repair, and at the present time, the only means we have for effecting this is to return the broken tapes to the manufacturer. Although each of these problems deserves continued attention, and still constitutes an annoyance, they have had only a minor effect on our reliability.

It might appear from the foregoing that there is little to recommend electronic computers, and the UNIVAC in particular. Despite the obstacles that stand in the way of reliability, the UNIVAC possesses many features that have more than tipped the balance in favor of reliability.

The mercury memory system has proven to be one of the most stable and reliable items in our operation. Except for some minor troubles causing poor heat control in some of the intermediate or "short tank" storage registers, we have never had trouble with the memory other than tube failures in recirculation loops.

Univac No. 2 is equipped with eight Uniservos, devices for transmitting information on magnetic tape to and from the computer. The Uniservos have proven to be essentially reliable. There have been many repetitive types of troubles: consistent failure of a specific transformer; periodic replacement of many components; and continuing need for adjustments. While not serious in themselves, these troubles do cause loss of time. Practical experience with the Uniservos has helped to increase their reliability. This has come about by a better understanding of their idiosyncrasies and methods for coping with them. One of the ways to reduce lost time owing to Uniservo difficulties is to secure at least one Uniservo over the number required for any expected type of computation. In this way a "spare" is present.

Prior to completion of the acceptance tests for the Uniservos, the reading heads over which the magnetic tape passes were wearing out at a rapid rate owing to the continuous friction of the metal tape against the head. This wear was reduced to practically zero by a modification that provides a thin spacing material between the head and the tape. The material does not reduce signal strength, but does reduce wear. This one modification has been a major factor in increasing computer reliability.

Probably the single most desirable characteristic of the UNIVAC, from the reliability standpoint, is the self-checking feature. Self-checking provides a built-in facility for servicing those parts of the computer with the duplicated circuits. When a trouble occurs in the duplicated part of the machine, it is rare that it takes place at exactly the same point in both paths of the duplicated circuits. Isolating trouble under these conditions may then consist of checking waveforms along the suspected path until a difference is observed.

Also, it has been generally agreed by all persons using the UNIVAC within the Air Force, that to our knowledge the UNIVAC has never introduced a mistake in computation. It has always stopped computation when an error was detected.

From the start of full-time operation until September 4, 1953, the computer had never been purposely shut down. As of that date, it was necessary to relinquish the 168-hour week in favor of a 120-hour week, owing to a decrease in the number of personnel. The machine was shut down at 12:45 a.m. on Saturday morning, September 5, and turned back on again for the first shift Monday at 12:15 a.m., September 7. The succeeding week produced a total of exactly 8 good hours of computation. The remainder of the time was used in servicing the computer. This first shut-down really took its toll. We removed bad tubes, especially marginal ones, by the score. The following weeks the situation improved, so that now, shutting down for a week-end has no serious effect. What caused the extremely bad week was the fact that marginal tubes never returned to their original condition, marginal though it was. The result was many tube failures. When these marginal tubes were removed, normal conditions returned. It is also interesting to note that the first five or six weeks after the heaters were turned off, the practice was to turn out all the lights in the UNIVAC area and conduct a search for open heater tubes. Since it was necessary to wait until the mercury tanks came up to temperature anyhow, two or three people could cover the entire machine in about thirty minutes, and in this way remove all cold tubes. In succeeding weeks, the quantity of cold tubes decreased continuously until at about the sixth week, the machine went right to work without any trouble shooting. Marginal checking rightly raises its head at this point. The Air Force has evolved a simple method for marginal checking. The preliminary work has been done, and we expect to have it installed soon.

Discussion of the technical problems involved in computer utilization is important, but there remains the one salient question: how much good time can one expect from a computer such as the UNIVAC in the hands of a crew such as ours? When we first put this computer on the air, we were determined to measure the time actually available for computation. We have made these measurements in the following manner: At the control panel of the UNIVAC are some switches labeled in accordance with the categories we wish to measure:

OFF—This covers week-ends or other time when the machine is not used.

INSTRUCTION PREPARATION—This is time used by the programmers to debug their routines.

PRODUCTION AND PROCESSING—This is time for actual computation.

PREVENTIVE SERVICING—Scheduled down time.

DOWN TIME—Unscheduled down time.

The items which constitute good time, i.e., time available for computation, are the sum of INSTRUCTION PREPARATION and PRODUCTION AND PROCESSING. The appropriate switch is thrown, in accordance with one of the foregoing categories, and there it stays until the situation changes. Any difficulty during a computation that consumes more than a few minutes is charged to down time.

This log has been kept as faithfully as possible, and any errors owing to the wrong switch being thrown are corrected when the reading is taken at the end of every eight-hour shift. With this explanation of how we arrived at these figures, I would now like to give them to you. For the approximate 18-month period from June 25, 1952 to December 4, 1953, the Air Force recorded:

GOOD TIME	61 per cent
PREVENTIVE SERVICING	20 per cent
DOWN TIME	19 per cent

For figures of this kind to have any meaning, they must be gathered over a long period of time. Certainly no less than a six-month period: preferably at least a year. Our highest monthly average was 74 per cent good time, the lowest 49 per cent. But on a day-in, day-out basis, the figures I've quoted are the ones actually compiled. Are they good? I can only answer this from the standpoint of the people I work for, and say that we've done an enormous amount of computation since the UNIVAC was installed. Could they be better? Most certainly, but only with additional people to allow much more preventive servicing than we've done. Given an 8-hour shift to do preventive servicing, your accomplishments are a function of available man hours. If six people could be put on an 8-hour shift, it stands to reason that we could do much more than the usual two persons we schedule. Available personnel are directly coupled to the amount of preventive servicing that can be done. Given these additional man-hours for this servicing, the amount of good time should be increased.

A question that may arise at this time, is what is the future outlook for more computational time for UNIVAC No. 2? Personally, I believe it must decrease; not very much, but some. We are just starting to get serious trouble from mechanical failures. I expect this to increase in the near future. Many items have run well past their specified life test. Most of these items, in the case of the UNIVAC, pertain to the input-output equipment, the Uniservos. They have always been rather ticklish propositions, and crochety old age isn't helping the situation any. Many of our relays are bound to give trouble as time goes on. This holds true for most all mechanical equipment. Also many of our undesirable crystal diodes of manufacturer type X are starting to go over the permissible tolerance in the forward direction. The result has been marginal trouble. With over half of the crystal diodes of this type, we can expect more failures. In fact, a series of troubles experienced in the past two months has been traced to crystal diodes which appear satisfactory under test conditions, but lead to intermittent difficulties when employed in the computer.

What all this adds up to is the fact that about three-fifths of our time has been put to computational use.

Discussion

Herman Lukoff (Remington-Rand): What reason do you have to believe that a G_2K tube short indication causes trouble in the computer?

Mr. Kopp: One of the best reasons that we have is that on replacing the tube the trouble disappears. I didn't mean to be facetious about this; actually it does represent a very serious problem. Mr. Lukoff has a very good point. What we have found is that the resistance of these G_2K shorts varies in value as read on a tube checker; going all the way from 10,000 ohms up to a million ohms. What we found we had to do was actually redesign the tube checker to indicate the value of these shorts. We do not know which tubes in which circuits can live with a G_2K short as low as 15,000 ohms. I would like to know this information. What we have done is worked on the assumption that if a tube shows a G_2K short and if our theory is correct that heat is causing it, we are bound to have more trouble later. Therefore, why not replace the tube right now, even if it isn't the real cause of the trouble you are experiencing? We have worked on this theory in discarding them.

C. T. Schaedel, Jr. (Consolidated Vultee Aircraft Corp.): Was fuse blowing the result of bad fuses or internal failure? Perhaps too close design to peak load conditions?

Mr. Kopp: I doubt if you mean, "Were the fuses bad from a manufacturing point of view?" I am inclined to discount this completely. What we do believe happened was that the fuses probably, in fact I am sure, are operated very close to the load conditions. In fact, in some cases the rated value was under the actual current through them. However, what we believe to be the real reason of grasshopper fuses blowing was a combination of possibly a transient plus the fact that we are having this heat trouble. We believe the combination of the two was actually the cause of the trouble but actually we can't put our finger on the specific details. I wish I knew more about it myself.

Lt. Com. H. N. Laden (Department of the Navy, Bureau of Ships): Please explain what parts of work time breakdown (good, down, maintenance) apply to the computer system and what parts to auxiliary equipment?

Mr. Kopp: If you mean by the auxiliary equipment the Unitypers and the Uniprinters, they have been discounted from these figures. If you include the Uniservos they have been included.

Com. Laden: What about card to tape converter?

Mr. Kopp: No information in regard to that. We have had it only a very short time.

L. Gallo (RCA): What is the status of G_1 when a short between cathode and G_2 exists in the 25L6GT?

Mr. Kopp: In the tubes that we have examined, and they are actually comparatively few relative to the number of tubes in which this kind of short has been experienced, there has been no apparent deposit of material on G_1 which we believe would cause any trouble. Now again, we may be far off base on this.

D. M. Rickerson (Department of Defense): Are all voltages automatically removed when one grasshopper fuse blows?

Mr. Kopp: Yes.

Mr. Rickerson: If so, why do clamping diodes blow?

Mr. Kopp: A very good question. It could be a transient condition. It might be that the diode is on the verge of possibly going anyhow. We really don't know.

Electron Tube and Crystal Diode Experience in Computing Equipment

J. A. GOETZ† AND H. J. GEISLER†

Introduction

THREE YEARS AGO, the first Conference on Electron Tubes for Computers was held in Atlantic City. At that meeting a paper of similar title was presented, outlining field experience with electron tubes employed in widely used IBM products. Data prepared at the time illustrated common causes and frequency of tube failure, as determined by laboratory analysis and statistics maintained on field replaced components. The program from which this data evolves is continuous, and since its inception in early 1949, has provided a substantial reduction in field tube failures. In some instances, improvement in machine performance has resulted from the introduction of newly developed tube types, and in others from corrections made in misapplication of existing tubes. It is the purpose of this paper to outline the program of component improvement phases of computer manufacturing.

Field Analysis System

IBM Customer Engineering field offices are the basic sources of defective components under study by our laboratory. Tubes from various environments and applications are sampled to obtain an adequate cross-section of over-all field usage. At present, this program encompasses 45 tubes and 3 diode types employed in IBM 101, 604, 605, 607, and 701 machines, and represents a total of 230,000 units. Fig. 1 shows the type and

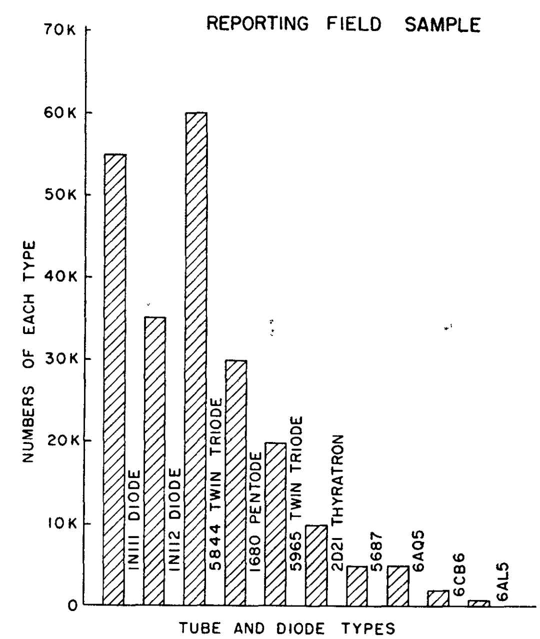


Fig. 1—Reporting field sample.

† IBM Electronics Laboratory, Poughkeepsie, N. Y.

quantity of each tube and diode included in the sample as of October 15, 1953. Co-ordinated development programs conducted between computer manufacturers and the tube and diode industry have resulted in the development of tube types 5844, 1684, and 5965 as well as crystal diode types 1N111 and 1N112.

The basic elements of the field program will now be discussed briefly. A number of extensive tube analysis programs have come into being for various ends since the initiation of the IBM system in 1949. An example of an excellent one aimed at improvement of military applications of tubes is that operated by Professor Walter Jones at Cornell University, under Signal Corps contract. All such systems in general retain fairly tight control of reporting machines and component removals, and most important of all, must have the co-operation of the human element involved at all stages from component replacement to final laboratory data filing.

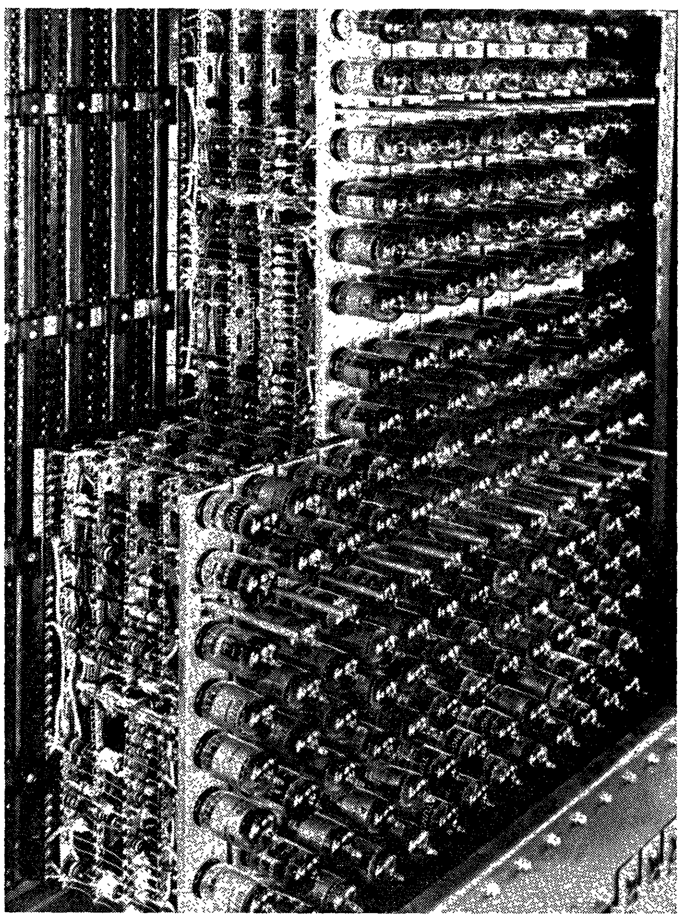
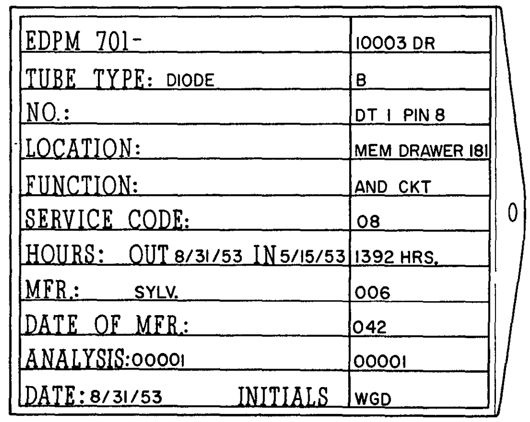


Fig. 2—Decalled tubes in 701 machine.

Our system requires that each tube in reporting equipment be marked directly with installation date, socket location at failure, and removal date. Additionally, these decals are numerically and alphabetically coded to identify the machine and office of origin for each tube. Typical multi-tube pluggable units bearing marked tubes in a reporting 701 machine are shown in Fig. 2. Where bulb temperatures or physical size makes the use of decals impractical, an enveloped stamped as in Fig. 3 is employed for field returns. The data shown on the envelope is from a typical crystal diode return. This so-called class B diode failed in a coincidence switching circuit for inability to recover properly (Analysis code 00001). The analysis laboratory uses the 5-digit code to report on mechanical, back resistance, stability, forward resistance, and recovery time characteristics. Field



EDPM 701-	10003 DR
TUBE TYPE: DIODE	B
NO.:	DT I PIN 8
LOCATION:	MEM DRAWER 181
FUNCTION:	AND CKT
SERVICE CODE:	08
HOURS: OUT 8/31/53 IN 5/15/53	1392 HRS.
MFR.: SYLV.	006
DATE OF MFR.:	042
ANALYSIS: 00001	00001
DATE: 8/31/53 INITIALS	WGD

Fig. 3—Typical return envelope.

failures in reporting machines are forwarded regularly to central office collection points, and are returned monthly to our Tube Analysis Laboratory in Poughkeepsie, N. Y. All tubes returned are tested for opens and shorts. Failures due to intermittent defects are uncovered by equipment capable of indicating faults lasting one μsec or more. Those failing to pass this test are examined internally to determine the exact cause of failure. The most common mechanical faults after machine installation are heater wire breaks, breakdown of heater-cathode insulation, and intermittent shorts. Electrical tests are made on a bridge in order to determine the actual shift in important characteristics during actual operation. These tests are made with fixed element voltages so that end-of-life limits can be established for circuit designers. In general, a drop of 40 per cent from the nominal zero-bias plate current value is considered a realistic end-of-life figure. Time-dependent changes in the E_b-I_b characteristic are frequently due to redistribution of cathode material and can alter the Class A plate current by much more than 40 per cent. Peak emission characteristics are measured on a pulse basis to avoid destructive testing. The results of these investigations, including data on machine and circuit of origin, total operating hours, source and date of manufacture, are punched into standard IBM cards for statistical purposes and future reference. The card shown in Fig. 4 contains all necessary information for future evaluation and is the basis of data presented in this paper. A complete analysis code has been developed for both tubes and diodes in order to accurately record the various faults. This code lists 200 different types of

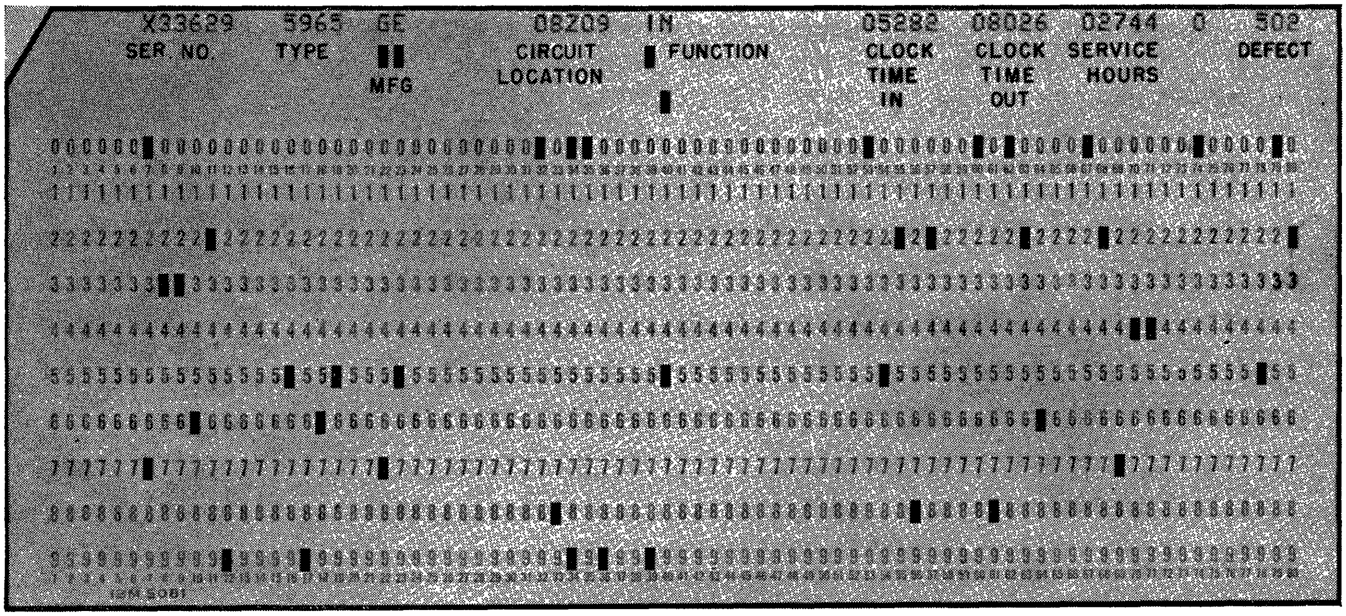


Fig. 4—IBM punched card for tube records.

electrical and mechanical failures. The information obtained from the cards enables the design engineers to improve tube utilization as well as indicate the course for development of new tube and diode types.

Program Results

The results of this program will be presented in graphical form. Fig. 5 shows typical rate-of-replacement curves for components used in IBM computing equipment. Curves apply to individual installations and contain all tube returns, including those resulting from liberal tube replacement during trouble-shooting periods.

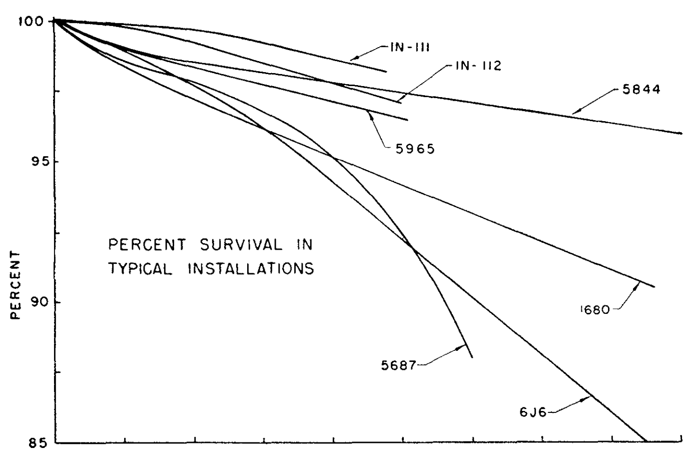


Fig. 5—Per cent survival in typical installations.

The 5687 appears to be an attractive tube for power cathode follower and inverter applications. However, recent field returns from our first 701 installation indicate that heater-cathode difficulties become increasingly apparent after 4000 hours. It is believed that the high operating temperature of the cathode contributes to the large number of tube failures.

A tube primarily employed in the 604 for switching operations is the 1680. Heater burnout is largely responsible for the failure of these tubes and has been traced to the use of nontungsten heaters.

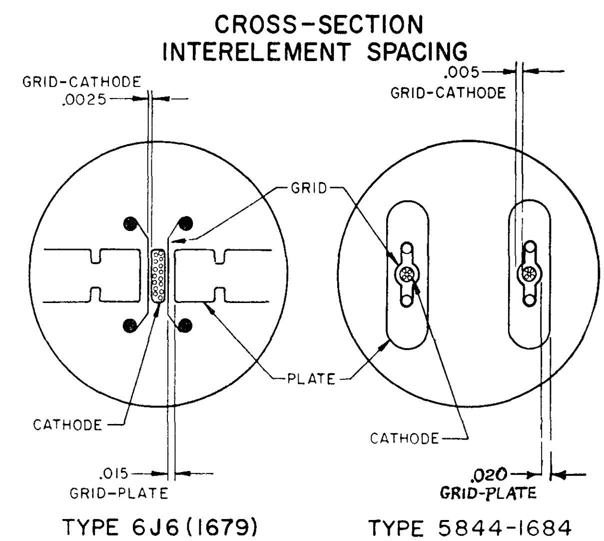


Fig. 6—Comparison of interelement spacing in 6J6 and 5844.

The 5844 is a long life computer tube used for trigger and inverter operation. This tube has replaced the 6J6 in most IBM field installations. Fig. 6 compares the internal structure and dimensions of the 5844 and 6J6. The close spacing and funnel-like structure of the latter tube increases the possibility of lint shorts, sectional instabilities, cathode deterioration, and inter-element leakage with life. It was the only available miniature double triode at the time of the original 604 machine development, and has provided tolerable performance for several years by reason of computer tube environment.

Fig. 7 illustrates the over-all failure rate for type 5844. The histogram shown at the top of the figure shows the number of tubes used for sampling purposes at various times; i.e., in Fig. 7 data is available for 7,500 hours on 5,000 tubes. Approximately 65 per cent of the defects were attributed to low plate current, 23 per cent to mechanical faults and 12 per cent to a shift in cut-off characteristics. These percentages are based on over-all tube failures for 8,000 hours. Mechanical defects are responsible for a very large portion of the initial failures, while plate current considerations account for most of the rejects after 1,000 hours.

TYPE 5844 AND 1684

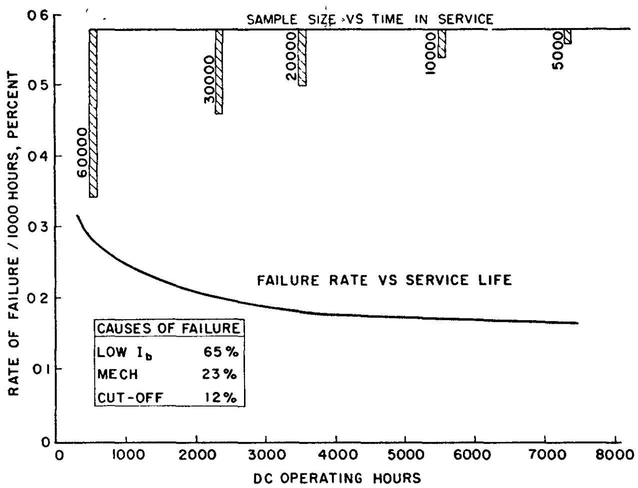


Fig. 7—Failure rate vs service life for tube types 5844 and 1684.

Fig. 8 illustrates over-all failure rate as a function of service life for type 5965, a 9-pin miniature double triode employed in the 701 EDPM. Low plate current accounted for 59 per cent of the defective tubes, gas for 20 per cent, and mechanical failure for 9.5 per cent.

TUBE TYPE 5965

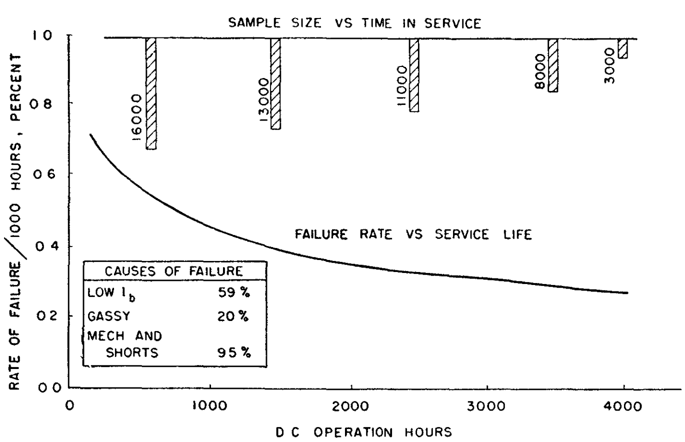


Fig. 8—Failure rate vs service life for tube type 5965.

This type is used widely for cathode follower and trigger circuits in high-speed computing equipment, and appears equal in performance to the type 5844 used in earlier applications. Interelement spacings are similar for both types, although the 5965 offers a sharper cut-off characteristic and increased plate dissipation ratings.

Fig. 9 illustrates computer experience with germanium diodes, types 1N111 and 1N112. Developed in conjunction with IBM needs, by the industry, these diodes are tested for forward and back resistance, recovery, drift, hysteresis, and flutter. This procedure was outlined in an earlier paper on diode testing by Mr. D. J. Crawford and Mr. H. F. Heath of IBM in March 1952. Many of the important characteristics are dependent on temperature and moisture content. The increase of failures with time is believed due to the fact that complex temperature and humidity cycling tests are not presently conducted on a 100 per cent basis.

CRYSTAL DIODE TYPES, IN-III AND IN-II2

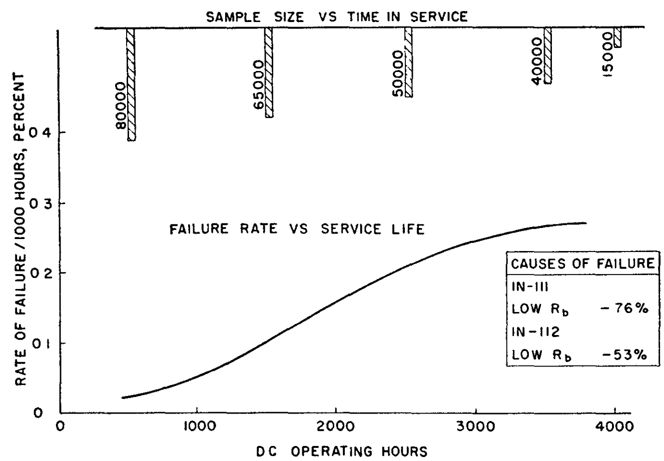


Fig. 9—Failure rate vs service life for crystal diode types 1N111 and 1N112.

Operating Conditions

One of the important factors in the study of tube and diode life is the environment in which the components operate. In computer design, many of the destructive forces contributing to tube and diode failures can be minimized. For instance, heater cycling is a well-known cause of heater burnouts. In computers, voltage is seldom interrupted, and when applied, is brought up slowly to ± 5 per cent of the recommended value. In cathode follower circuits the heater is dc biased to avoid destructive electrolysis effects in the heater-cathode insulation. Air is forced around units containing tubes, limiting bulb temperatures on such types as the 5965 and the 5844 to a maximum of approximately 60 degrees C. This reduces the possibility of gas evolution and cathode poisoning during operation. Tubes are operated under conservative ratings on plate dissipation and interelement voltages that have been established by extended life tests. Vibration or shock is not expected following installation, and equipment is generally shipped with precautions observed as to handling. This allows tube construction with a minimum need for mica and its well known poisoning effects on the cathode. Warping of the tube elements caused by tight mica construction is also reduced.

Incoming Inspection

All these environmental precautions must of course be supported by good quality tubes. The failure of a single unit out of several thousand components can cause a computer to make an error. So, intensive inspection procedures are followed for tubes and diodes used for both original installation and field replacements.

Mechanical inspection criteria for tubes include such items as inferior getter flash, cathode coating flaking, weld particles, broken micas, and poor glass quality. While these conditions may not always show up initially as intermittent shorts, noise, gas, or microphonics rejects, they nevertheless represent a hazard where a large population of tubes must continue to function at very low failure rates to insure good machine operation.

Besides electrical and mechanical tests, random samples of production lots are periodically placed on life test to investigate short life conditions that might appear due to material or processing changes in production. This precaution has been necessary because our customer engineers must have tube replacements that are trustworthy in order to avoid a random distribution of tubes destined for short life. After 500 hours operation under rated conditions, 90 per cent of the readings involving inter-element leakage, cathode emission, stability, and contact potential must be within established limits before the lot can be accepted.

Circuit Operation

Life tests of time-dependent characteristics must be evaluated so that circuit designers can assign realistic end-of-life limits to tubes and diodes. These tests develop data on such items as changes in inter-element leakage, increases in cathode interface impedance, changes in contact potential, normal cathode emission decay rate in tubes, and the development of hysteresis, drift, and other dynamic instabilities in germanium diodes. The presence and degree of such processes is particularly important in new machine development and can frequently be used to predict field performance.

For example, in some cathode-follower circuits, machine failures have been traced to the appearance of grid current and a change in grid voltage level during life. Cathode-follower tubes will gradually advance into positive grid current regions in order to maintain a given plate current throughout life, and this results in a loading down of preceding circuits. Also, where high resistance dividers are employed in the grid circuits of class A amplifiers or cathode followers, the flow of a few microamperes of grid current at -0.5 volt bias can cause great deviation from the originally designed operating point. Variation in contact potential, initially and during life, can sufficiently alter cathode-follower levels to require level setter circuits to be incorporated in the equipment. In triggers, inter-element leakage paths caused by redistribution of volatile metals can become great enough to cause machine failures. The development of cathode interface as a cause for trigger failures is also well known. Protection against any of these time-dependent characteristics can only be obtained by periodic life tests under various computer circuit conditions.

Extended life tests are also conducted to evaluate peak cathode-emission capabilities of tubes destined for duties involving blocking oscillator and high-speed switching-circuit functions. Most important of all, it is necessary to determine the degree of safety factor in various voltage, current, and dissipation ratings applied to tubes whose warranty periods normally expire within one year of the date of sale. Few tubes have been examined by tube manufacturers for all pertinent characteristics over extended periods of time. This applies particularly to heater cathode voltage ratings which are of considerable importance in cathode-follower design. Unfortunately many of these ratings contain no safety factor and because accelerated testing procedures are inadequate little data is available for 10,000 hours of operation.

New Tubes and Diodes

Computer circuit designers are constantly looking for tubes that can supply large amounts of plate current at low plate voltages in the overdriven case and which will cut-off at relatively small negative grid voltages. To do this without increasing the hazard of intermittent shorts becomes a difficult problem for the tube manufacturer. The Tube Analysis Laboratory is constantly involved in evaluating new tube developments in the light of these basic requirements.

The picture for crystal diodes is complicated by the fact that what a computer circuit designer really wants is low forward resistance and high back resistance with rapid recovery at elevated operating temperatures. This is found only at the edge of normal diode distribution curves. However, compromises on instabilities due to poor contact and poor moisture proofing cannot be tolerated and life tests must be maintained to insure satisfactory machine operation and to determine the reliability of ratings.

Application Memoranda

All tube and diode experience from the field as well as the laboratories is gathered in the form of application memoranda. A typical portion of a 13-page memoranda on 2D21 thyratrons is shown below.

Project #5000-5021-109-0009[1]

3.1.2 Importance of Ratings (Continued)
1. heater voltage
2. minimum plate voltage
3. negative transients, commutation factor
4. cathode current
5. amps/μsec.

3.2 Coincident Switching Notes

3.2.1 Since thermionic emission causes loss of signal at positively driven control or shield grids, the range of decoupling resistor values that will limit ionization current without loss of signal is limited. Preionization currents are particularly important when coincident switching employing 2 grids is attempted at low plate volatges. Under these conditions, a priming gate voltage of 25 volts and a conduction current of 10 ma is required. The shield grid has a larger capacity and draws more current than the control grid; it should be driven from a low impedance source.

3.2.2 Shield currents up to 100 ma have been drawn without loss of plate control by grid 1.

3.2.3 It has been determined that reliable firing requires the shield grid voltage to be stable within 10 per cent of priming amplitude at least 10μ secs. before the control grid is required to fire the tube.

[1] *Preliminary Report, November 9, 1953, Sheet 5 of 13*

3.2.4 At low plate currents it is important to:
(a) Decouple the plasma oscillations from common supplies and signal bases.
(b) Keep the cathode from generating self bias.

3.2.5 Where inductive loads are used, resistance or diode dampening techniques should be employed. This will prevent:
(a) Loss of control due to cold cathode discharge.
(b) Reduction in tube life due to back conduction and gas cleanup by positive ion bombardment.
(c) Cross talk between circuits and spurious firing.

3.3 Pulse Application Notes

Attention is particularly directed in such memoranda to electrical characteristics actually controlled by the manufacturer for a given tube. The circuit designer is thus forewarned to avoid dependence upon noncontrolled and unrealistic characteristics.

Conclusions

Although great improvement has been made, tubes and diodes are still the major source of unreliable operation. We believe that the elements described in this program have made it possible to visualize a more desirable computer tube. Continuing co-operation between manufacturer and user will assure the further development of reliable computer components.

Discussion

J. J. Scanlon (Bell Telephone Laboratories): What is the nature of intermittents of microsecond duration?

Mr. Geisler: Although our equipment can detect these very short duration intermittents, they are very rare. Most of them appear in the millisecond or tenths of millisecond range. I might add that the tube is rotated while being tapped 10 times with two tappers in quadrature at approximately 50g., for ½ millisecond. It is difficult to describe a mechanism for the microsecond duration paths that have been measured.

W. L. Martin (Telecomputing Corp.): What are some other common causes of diode failure that you have observed besides low back resistance?

Mr. Geisler: The IN 112 crystal diode will frequently fail for its inability to recover—to 200 K back resistance—in the prescribed 1 microsecond. For this test 35 volts is applied thru 2 K in the back direction one tenth microsecond after forward conduction of about 30 mils.

J. Howard Wright (National Bureau of Standards): Can you comment on the envelope temperature of the 5687 in service? If it is high, do you reduce filament input power accordingly?

Mr. Geisler: We do not put decals on the 5687 because we are running near full dissipation and even with forced air cooling, the bulb temperature is around 120 degrees C. We do not reduce the cathode heater power on the 5687 in the 701 installation.

J. L. Raymond (Northrop Aircraft Corp.): What is the nearest to zero grid-cathode bias voltage one can use without contact potential effects?

Mr. Geisler: This depends upon the tube type; approximately 1 to 2 volts negative should be adequate for most of the computer types.

D. Davies (National Physical Laboratory): Do you experience much difficulty with your 2D21 thyratron applications?

Mr. Geisler: We experience exceptionally low failure rates for our 2D21 applications. However it is very important that the maximum and minimum ratings include any transient conditions that may exist.

Reliability and Characteristics of the Illiac Electrostatic Memory*

J. M. WIER†

Introduction

THE ILLIAC MEMORY was completed in the spring of 1952. It was tested during the early part of the summer and in August was attached to the previously completed arithmetic unit. About three weeks were needed to complete the physical work necessary for this and to eliminate all unsoldered joints, misconnections, etc. Detailed records were kept after this time. This log begins on September 3 and covers every "on" period in detail. Later entries are more specific due to the added experience gained in isolating difficulties.

The memory is of the Williams type. It is a parallel device with 40 words of 1,024 bits each. The storage tube which has been used is the 3KP1. It is operated with 2,000 volts accelerating potential. The regeneration amplifier has a gain of about 70,000 to a positive pulse and is a very straightforward RC coupled type. The nominal bandwidth is 60 kc to 500 kc. The coupling time constants are 5 microseconds each. The period of a regeneration cycle, that is, the time to regenerate one 40-bit word is 19 μsec.

Philosophy of Operation

During the first year of its operation a number of changes and improvements were made. Since it was recognized at the outset that this would be done, it was felt desirable to set aside time for these changes as well as time for periodic checks of machine operation. The

* This work was supported in part by the Office of Naval Research.
† University of Illinois, Urbana, Ill.

8:00 AM to 12:00 noon interval was chosen for this. The running period is considered to be from 12 noon to 12 midnight, five days per week. The Illiac is also on from 8:00 AM to 12 noon and occasionally at other odd hours, but these times, although completely logged, are not considered as running time, even though some computational work may be done during them.

This report is limited to the running period. All time in this interval is accounted for including any unscheduled maintenance required in the interval. All known errors in this interval are recorded. For this paper an error is considered to be a known and detected failure. If two failures occur in five minutes, even though the second occurs while in the process of isolating the cause of the first, this will be treated as an additional error. Such errors make up the majority of short interval failures.

A number of tests are run at regular intervals to determine the operating condition of the machine. This is done about 5:00 to 5:30 during the running period. Special tests of the memory alone include a read-around ratio test for testing address interactions and flaw tests to determine the existence of nonstoring spots at addresses used. Since the Illiac never has possessed a complete set of flaw-free tubes, it has been necessary to operate on cathode-ray tubes which contain nonstoring spots. A relatively dc-stable address generator used in combination with small trimming magnets placed near the storage surface of tubes known to contain flaws has been partly successful in keeping storage points off of nonstoring spots on the cathode-ray-tube screen.

A further and more exhaustive test is performed by using a so-called leapfrog test.[1] This test is designed to detect errors of all types which may occur, with the exception of errors in the printing of results. Due to a closed loop through the output punch, even the input and output are checked. This code has the advantage of isolating memory errors from other types and giving their memory locations. It is through its use that this report is able to give the cause of a given error in most cases. There are, of course, errors that occur at other times which are not so well defined and these often are not specified as to cause.

One characteristic of the Illiac causes the existence of undetected errors to be much less likely than they might at first sight seem to be. Since the instruction code of the Illiac is not completely decoded, there are a number of possible arrangements of the instruction digits which are meaningless and upon whose appearance the machine will stop. Probably the most useful of these is the order which is represented by all zeros. Since the memory is cleared to all zeros before a problem is run, if something goes wrong with a code, it is usually shortly in one of two conditions, looping or stopped in the all zero region of the memory. It is then evident that such failures are easily detected and it often turns out to be a simple matter to locate the failure which has caused this behavior. Of course, random and very infrequent failures in the arithmetic operations are more difficult to isolate although it has been found in general that such errors soon make themselves more and more evident and do not go long undetected.

Since this report covers the memory it should be pointed out that memories in general have one property which makes the determination of an error quite simple. Unless they fail completely, they tend to remember their errors.

[1] D. J. Wheeler and J. E. Robertson, "Diagnostic programs for the Illiac," PROC. I.R.E. vol. 41, pp. 1320–1325; October, 1953.

Collection of and Pertinent Comments about the Statistics

As has been mentioned this report covers the noon to midnight period for five days per week. Since the machine is turned off completely every night there is a larger probability of machine failure in the period 8:00 AM to 12:00 noon than in any other four-hour period. This is because of the disturbances caused by the turn-off-turn-on process. The error rate is in general largest then. This period is used to attempt to place the machine in such a state by noon as to run error-free the remainder of the time. The degree of success attained in this may be seen from the statistics which follow.

There are a few disturbing factors which make this report slightly pessimistic. Certain of the running periods were immediately after major engineering changes and required troubleshooting to iron out errors caused by such prosaic things as unsoldered connections or, occasionally, logical errors. There are three such periods rather carefully distinguished, the first one being in the first two weeks after the final assembly of the machine; the second one after the complete replacement of the adder by a more efficient type; and the last one by an operational change in the memory which improved the read-around-ratio figure. The adder replacement probably does not affect the memory figures as much as the others.

The figures given encompass only errors which were reasonably assignable to the memory. It is possible that a few others may have been due to the memory, since about 80 errors in this period could not be traced.

Statistics

These statistics cover a period starting on September 3, 1952 and ending August 8, 1953. During this time the Illiac logged 4,100 dc hours. This report covers 2,976 hours of this, falling in 12 noon to 12 midnight interval.

During this time a total of 479 known machine errors occurred. Of these 179 were directly assigned to the memory and 80 were not traced as to cause. Thus the memory caused a minimum of 37.4 per cent of the total number of errors and a maximum of 54 per cent of them.

The errors attributable to the memory were investigated and it was found that 123.2 hours of computing time were lost due to the memory. This lost time includes all the time lost due to running the problem terminated by the error, the time necessary to make repairs, and the time to make sure the trouble had been cured. The latter process usually consists of the running of test programs covering the problem at hand. This lost time is 4 per cent of the total period.

Since there were 179 memory errors in 2,976 hours and 80 errors not traced, mean free time between memory errors is 16.6 hours at best and 11.5 hours at worst.

Although mean free time figures are instructive, it was felt to be desirable to analyze the errors more completely. This was done by making a table of all known memory errors in this interval, noting the time of each error. Then the time between successive errors was found, considering each 12-hour period to be immediately succeeded by the next 12-hour period. These differences then indicated the different times between memory errors. These differences were grouped by length of error-free run and counted. Fig. 1 shows the results of that process. Note that the time intervals are growing exponentially larger to the right. All curves are normalized to the total number of memory errors. It should be noted that the 8–16 hour group is the largest.

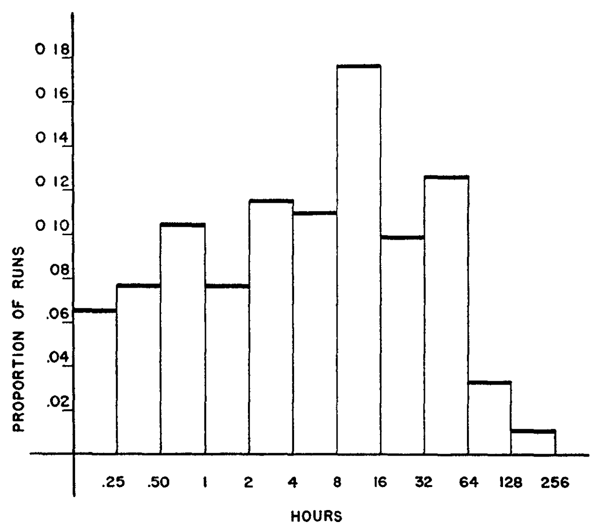


Fig. 1—Length of run distribution.

This is probably due to the disturbing influence of turning the machine off between each successive set of 12 hours. The error-free intervals from 0 to 1 hour are largely in the unscheduled maintenance periods.

Since the ordinates are normalized the height of any given group bar is the part of the total number of memory errors in each group. In order to find the probability that a given memory run will be longer than some time, we may add all the groups to the right of that abscissa. The results of doing this are shown in Fig. 2. Thus if one takes a time immediately after an error occurs and wishes to know the probability of running a problem of length H hours, this may be found by going to the abscissa H and reading off the probability on the axis of ordinates. It will be noted that the point at which the probability is $\frac{1}{2}$ is somewhere in the 4–8 period, being nearer the latter. The fact that this does not agree with the 11 to 16 hour mean free time stems from the large number of closely spaced errors in the unscheduled maintenance period.

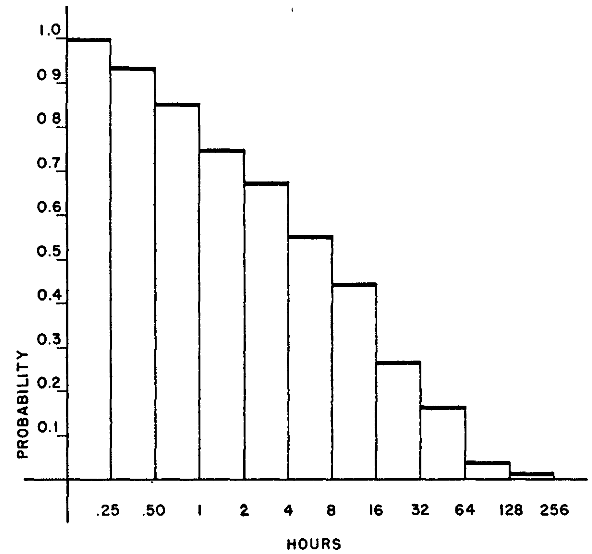


Fig. 2—Probability of a given run length.

Causes for These Errors

In order to determine what caused these failures, each error was catalogued as to a type of error. Table I shows how these errors are distributed:

TABLE I

Causes for Errors in the Memory

Cause	Number of Errors
Flaws	81
Unknown	71
CRT Intensity Drift	10
Poor CRT Focus	7
Address Generator	6
Changes Logic Circuits	2
Noisy Regeneration Amplifier	1
Poorer than Stated Read-Around Ratio	1

It will be noted that there is a large contingent of unknown errors. Most of these were accumulated in the early history of the machine when the operators were not so expert in finding or analyzing the errors. Also, the leapfrog test code in use then was not so completely able to isolate the trouble. It is evident that by far the commonest error involves flaws on the cathode-ray tube surface. As has been said, the Illiac operates in the presence of these flaws and from time to time components drift in the address generator causing the difficulty noted. Other troubles are few and are usually fairly simply repaired. It should be noted however, that the largest part of the trouble has been due to the cathode-ray tube, a component which has also had the most effort placed in the direction of its improvement by various manufacturers and computer groups. All of the tubes of an improved nature which were tested were found to be capable of eliminating the majority of these errors. These include sample lots of the IBM 79, the RCA C73376, and C73621 types.

After diagnosing these errors, various corrective actions were taken. The most frequent of these was to place a small magnet near the tube face to move the raster off the flaw. Although this is felt to be somewhat makeshift procedure, it has been reasonably successful in holding flaw difficulties down.

The original set of 40 cathode-ray tubes was selected from a set of 160 standard 3KP1's. About 50 per cent of this number were considered to be good enough for retention as possible storage tubes. The remainder were discarded either for the presence of more than two nonstoring flaws or for reasons of poor read-around ratio. Forty of the better tubes were placed in the memory for preliminary tests before connection to the Illiac. The remainder were used as replacements. Twenty-five additional tubes have been procured since then for replacement purposes.

During the period September 3, 1952, through August 8, 1953, it was necessary to replace 43 cathode-ray tubes. Table II shows the distribution of difficulties for which these changes were made. It is evident that the principal reasons for removal were flaws and low read-around ratio; these accounted for 30 of the 43 replacements. The read-around ratio removals were made to improve that quantity. Tube changes were made for reason of flaws only when two nonstoring or marginal flaws existed at storage points, at the same time.

TABLE II

CATHODE-RAY TUBE FAULTS

Fault	Number of Removal Caused by Fault
Flaws	16
Rar	14
Suspicion	8
Poor Focus	1
Shorted Elements	1
Loose Aquadag Connection	1
Open Filament	1
Remote Cutoff	1

Of the remaining 13 tubes, 8 were replaced on suspicion that they were causing difficulty, but without very rigorous evidence. It is quite possible that some of these replacements were unnecessary.

The chassis housing the regeneration amplifier and circuitry necessary for the logical operations needed to write in and read out of each position was a relatively trouble-free device. As may be seen from Table III, it was found necessary to remove 16 chassis for repair in the September 3, 1952 through August 8, 1953 period. Those removed on suspicion were generally found to have poor amplifier tubes which resulted in an occasional noise spike coming out which resulted in an error. In this sense they are indistinguishable in symptoms from those labeled noisy amplifier and microphonic. Only one other failure was found in this category and that was a leaky bypass condenser.

Those errors listed in the logical section were errors which completely prevented storage. Four of them were caused by open filaments and one by a grid-to-plate short.

Thus the failures in the memory were almost exclusively tube failures.

TABLE III

REGENERATION CHASSIS FAULTS

Reason for Removal	Number of Chassis out for this Reason
Suspicion	8
Logical Section	5
Microphonic	2
Noisy Amplifier	1

THE ADDRESS GENERATOR

The address generator forms the analog equivalent of any given digital address and deflects the cathode-ray-tube beam to that address. It caused six errors. Four of these were due to tubes with open filaments and the remainder were grid-cathode shorts.

In addition to those actions taken after failures had occurred, a considerable amount of preventive maintenance was done. This maintenance included tolerance tests of various supply voltages, hammer testing of circuits to locate microphonic tubes or circuits, and a complete test of all tubes except those in the regeneration chassis, at about 3,500 dc hours. This latter check resulted in the replacement of a relatively large proportion of the type 5687 tubes which are used extensively as drivers.

It should be mentioned that the set of ten pulsers and their output drivers caused no known trouble in this period although a number of tubes were replaced in them at the time of the wholesale tube check. Certain tubes in the address generator were also removed at this time without their having caused any known error.

READ-AROUND RATIO

One figure of merit for all Williams tube systems is the read-around ratio. Since there is some interaction between adjacent storage spots in the tube, too frequent consultation at some given address without regeneration of its neighbors tends to cause loss of information in these neighboring addresses. Of course, as the packing density increases, this interaction becomes greater. It appears that 1,024 storage spots per tube is the practical limit within a factor of two with present tubes used in a parallel manner.

In order to improve the read-around ratio figure, various efforts have been directed toward improving the cathode-ray tubes and the method of operation. The Illiac memory system, limited to standard cathode-ray tubes because of procurement difficulties, has used circuit improvement. R. Thorenson of SWAC proposed last February that better results might be achieved by sensing the presence of information at a later time in a typical output pulse.[2] A scheme using this information in a way slightly different from that proposed by Thorenson was devised for the Illiac. After tests in a test rack this was found to be sufficiently better than the normal Williams system to make installation in the Illiac worthwhile. As the changes were minor, this took only a weekend. The result was a net improvement by a factor of three to four in the read-around ratio.

The test for the read-around ratio is programmed and tests every point on the raster in the following way. The half of the raster to be tested is filled with 0's. 0's correspond to the longest beam on time with the Illiac system. The addresses surrounding the test address are filled first with 0's and then the test address is bombarded n times with 0's, (where n is the test read-around ratio number). The surrounding addresses are checked for failures and these are noted, if any. This is done twice at the address to avoid the possibility of stray regenerations making the test less severe. Then the surrounding addresses are filled in with 1's and the bombardment and checking takes place as before. Failures are printed out with suitable indicating information to give the test number n, the direction of failure, (from 0 to 1 or 1 to 0), and the address and tube failing. This test for failures both from 1's to 0's and 0's to 1's is unnecessary with the ordinary Williams system operation since failures occur only in one direction. With the Illiac system, a more balanced one, both failures occur with similar frequencies.[3] The point of marginal failure is 3 to 4 times better than the normal Williams system. This system has been used on the Illiac since March 1953 and since then the read-around ratio value from the standpoint of the programmer (at which no failures occur on any tube at any address) has varied from 60 to 80 depending on the condition of the tubes in use. This has been enough to free the coders of excessive care in coding against read-around ratio. It is possible to cause failures if sufficiently short, fast loops are used.

[2] R. Thorenson, "An Improved Cathode-Ray Tube Storage System," *National Bureau of Standards Report 2275*; February 6, 1953.

[3] J. M. Wier, "The Illiac Memory System," *Proc. of Symposium on Large Scale Digital Computing Mach.*, Argonne Nat'l. Lab.; August, 1953.

Summary

The following statistics are of greatest significance. Machine time lost due to memory failures totals 4 per cent of 2,976 hours of running time. A total of 179 errors were directly attributable to the memory for a mean free running time of about 16.5 hours. Starting after an arbitrary failure of the memory, there is a probability to $\frac{1}{2}$ that a continuous error-free run of at least 4 to 8 hours will follow, the true figure being somewhat closer to 8 hours. The read-around ratio is a minimum of 60 to 80 on all tubes.

In the September 3, 1952, to August 8, 1953 interval, including time not called running time, all cathode-ray tube replacements totaled 43. It was necessary to repair 16 regeneration amplifier chassis in this same time. Other difficulties were few and simply repaired.

Further Improvements in the System

As the principal cause of failures has been the cathode-ray tubes, a complete set of the RCA type C73621 has been ordered and partially received. In order to eliminate the other errors a regular system of checking vacuum tubes has been instituted to catch tube failures before they cause machine errors

A number of experiments using more elaborate discriminating methods have been tested to make the system more independent of operating variations. The most successful of these involves an integration sensing method which integrates the magnitude of difference between the desired signal and the signal present over the whole meaningful output pulse. This system is very insensitive to noise in the system but involves somewhat more circuitry than can be added to the present regeneration chassis.

It is felt that the system has reached a state of reliability compatible with the error rates of the rest of the machine with the exception of the state of the cathode-ray tube. Further improvements, while desirable, cannot add too effectively to the stability of the system as a whole, since the present Williams memory is nearly as reliable as the average for the remainder of the machine, judged on a per tube basis.

Discussion

L. E. Kanter (IBM Corporation): For every reference to the memory is there a fixed regeneration cycle period or how does the time allotted to execution of instructions compare with time allotted to regeneration by the ILLIAC?

Mr. Wier: The memory is basically driven by a clock which has a period of 19 microseconds. In a sense the arithmetic unit is a slave of the memory. The arithmetic unit waits until it reaches the time in the memory cycle at which it can acquire access to the memory. However, it is so arranged that it is impossible to require that the memory give forth two numbers in two successive regeneration periods. For this reason the upper value on the read-around-ratio to which it is possible to subject the machine is limited because every other period at the very least has to be a regeneration cycle.

J. L. Cochran (Defense Department): How is heater voltage applied to ILLIAC memory tubes? That is, the length of time from zero to full voltage?

Mr. Wier: During the period covered by this report, the voltage was applied abruptly. I will not defend this, however we have not had an excessive amount of trouble. Currently the voltage is applied as it has been all the rest of the time on the remainder of the machine. The voltage is applied through a large and nonlinear resistance which consists of a series of heater elements. The voltage starts out at essentially zero, builds up slowly to about 3 volts and then these heater elements are shorted out and the voltage goes abruptly to the value at which the filaments operate, namely, 6 v.

S. Greenwald (National Bureau of Standards): Would you give some detail on magnets used to avoid blemishes, such as shape, strength and location? What is the effect on CRT focus?

Mr. Wier: I am in a very bad position to answer this because the magnets which we are using are of the "friendly dog" type and I do not have detailed characteristics of these magnets. I think the only comment which might be made is that they are weak magnets and they are not standardized. As to their effect on focus we have found that there has been no lessening of any quantity which would normally be a function of focus such as the read-around-ratio unless the magnet is placed closer than about $\frac{3}{4}$ inch from the face of the tube.

Oliver Whitby (Stanford Research Institute): After the machine is first turned on in the morning, how long is it before the incidence of failures settles down to the average rate noted for afternoon operation?

Mr. Wier: The reason that I have avoided the mornings is because of the fact that we have made a large number of engineering changes and this process is still going on. This is not because the machine abruptly becomes a pile of junk at eight o'clock in the morning. When the machine is turned on, it may reasonably be expected to run any one of the test programs 5 minutes after it has been turned on. It usually takes this long to achieve a stable charge on the face of the cathode-ray tubes. I would say that probably the incidence of failure during this time is slightly higher than it would be during the afternoon, even after 5 minutes, but at the end of an hour the operation is substantially identical. The only reason that I have avoided the morning period is that I wished to avoid classifying periods which were not used for preventive maintenance but used only for changing the equipment.

Electron Tube Performance in Some Typical Military Environments

D. W. SHARP†

THE INFORMATION presented in this paper is concerned with the performance of electron tubes in various types of military equipment. While these equipments do not include computers, some of them resemble computer applications to such a degree that I believe the tube data will be of interest, and perhaps helpful, to this audience.

The data have been selected from the mass of information assembled by Aeronautical Radio, Inc., through field surveillance of equipments in actual military use. We are conducting this surveillance program for all three military services under a contract issued through the Navy Department. The program is being carried out at various bases selected by the military to provide a good cross section of all electronic tube applications in all types of environment.

The applications include communications equipments ranging in complexity from the small handy-talkie to high-powered fixed installations, airborne and ground radar, sonar, gunfire, and missile control systems, IFF, ECM, and others. Among the types of environment represented are man-pack sets, wheeled and track vehicles, propeller-driven aircraft, jet fighters and bombers, ships of the U. S. Navy, and fixed-station, land-based equipments. It is felt that this presents a fairly complete coverage of military tube applications as they exist today.

Before beginning the discussion of the data, let me outline briefly our methods of collecting tubes and data in the field and compiling the information for analysis.

Methods of Collecting Data

When a base is selected by the military for surveillance, an ARINC field office is set up at that base. The military electronics personnel are instructed to do two things: First, to place all failed or rejected tubes in a printed envelope provided by ARINC; and, second, to fill in the information requested on the envelope. This information includes the unit of equipment and number of the socket from which the tube was removed, the tube type, the reason for removal, and pertinent facts concerning the condition of the equipment at the time of removal. If the removed tube is of a type that is too large for insertion into the envelope, the envelope is attached to the outside of the tube.

The tubes and envelopes are deposited in containers conveniently placed at repair stations, from which they are periodically collected by ARINC field representatives. All tubes thus collected are marked with a serial

† Aeronautical Radio, Inc., Military Contract Div., Washington, D. C.

number identifying the tube and the week of collection. Receiving-type tubes are tested on a Hickok 539A tube tester, and the results of the test, together with the information on the collection envelope, are recorded on a weekly tabulation sheet.

Once a week, the field office ships the collected tubes and the tabulation sheets to ARINC headquarters in Washington. Here, a sample of the tubes are retested on laboratory tube testers. This is done as a means of providing additional information and, also, as a continuing check on the accuracy of the field testers. An IBM card is punched on each tube, making a permanent record of the tube's history, and the tubes are then forwarded to Cornell University for further testing. The results of this final testing are made available to ARINC and included in the permanent record on each tube.

Tubes collected in the manner I have just described are referred to as "semi-controlled." This term is used to indicate that we do not have complete information on the tubes—that is, we do not know the hours of service, nor do we know the exact condition at installation, other than the fact that the tubes were new when installed. To obtain these types of information, we conduct what we refer to as "controlled tests." The distinguishing feature of a controlled test is that the tubes are installed by, or under the supervision of, ARINC personnel, and complete information is obtained. Each tube is individually numbered by means of a decal, the tube is pre-tested, and the test readings are recorded. Operation time is kept on each tube until it fails, or until the test is terminated. Failed and terminated tubes are retested and the results recorded. Complete data on each of these tubes are punched on IBM cards.

The information on tube performance which I shall now present is from two of the military bases in our surveillance program. One is a land-based communications station; the other is a naval base.

Land-Based Fixed-Communications System

At the first base to be discussed, the equipment includes high-powered transmitters, receivers, microwave relay installations, and terminal equipment such as teletype receiving and transmitting units. The equipments are all rack-mounted, with space and weight considerations secondary to accessibility and ease of testing. For the most part, they are located in air-conditioned rooms, so that high ambient room temperatures do not occur. The equipment operates 24 hours per day except for periodic maintenance; therefore, cycling is not a factor in tube wear-out. The power source is well regulated, eliminating surges to the units. This base was selected for discussion here because it was believed that the electronic applications would more closely parallel computer application than those at other surveillance bases.

Maintenance procedures at this base differ somewhat for different equipments and are based on experience with the equipment performance.

Generally speaking, equipments are tested in the racks where they normally operate. The receivers, for the most part, are given a daily performance check. If no trouble is indicated, no further handling of the units occurs except for a quarterly or semi-annual testing of all tubes, at which time tubes are replaced if they fall below a predetermined G_m value for amplifier types and an emission value for diodes. The perodic testing is performed as often as experience indicates is desirable, but no more often than necessary to hold emergency failures of equipment to a minimum.

Maintenance on the teletype terminal equipment differs in that performance is tested frequently at reduced power supply voltage. If the units function normally at the reduced voltage, no other test is made and tubes are not tested or replaced on any periodic basis.

It is felt that tubes in these equipments are given excellent chances of survival. Factors usually considered detrimental to tube life—such as vibration and shock, cycling, handling, and frequent testing of tubes on tube testers, high operating bulb temperatures, temporary or permanent over-voltages, and compact weight-saving construction—either are held to a minimum or are nonexistent.

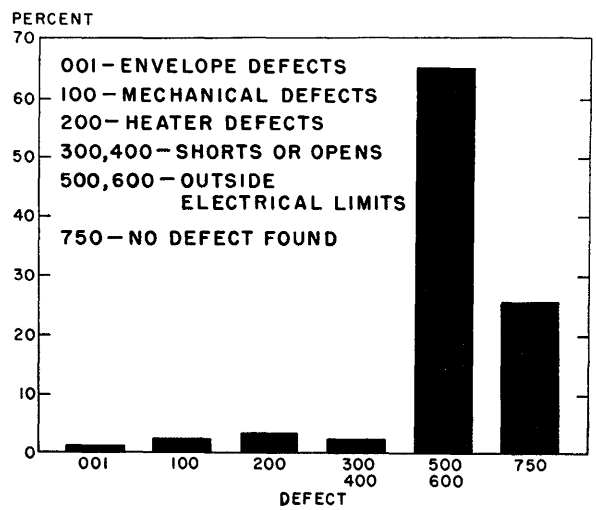


Fig. 1—Defect distribution of tubes removed from land-based communications equipment.

A sample of 2,343 tubes returned from this base have a defect distribution as shown in Fig. 1. The defect findings on which the distribution is based were obtained, in the case of receiving-type tubes, by testing each tube on our field tester; for transmitting types, the defect findings represent the judgment of the military maintenance personnel. The tubes represented here do not include any of the military improved types manufactured under Navy and Air Force programs.

The first bar on the left side of Fig. 1 represents tubes with broken glass. It is not known that this glass breakage occurred during normal operation; some of it may have occurred after the tubes were removed from the equipment and before retesting in our laboratories.

The second bar represents visually observed mechanical failures. These faults include loose plate caps, loose or broken bases, and broken welds and supports within the tube.

The third bar represents heater and filament failures, exclusive of heater-cathode shorts.

The fourth bar indicates the percentage of returns having shorted or open elements within the tube, including heater-cathode shorts.

These first four categories may be grouped under the broad general classification of mechanical failures. This type of failure is considered catastrophic—that is, the tube functions up to a point, then fails abruptly. From the base in question, the total of these mechanical faults represents only 9 per cent of the aggregate returns.

The fifth bar, as can be seen, accounts for a majority of the tube returns. The tubes in this defect category are considered electrical failures, meaning that they failed to meet the lower-limit test on some parameter by which tubes are judged. Low transconductance is the predominant defect of the tubes in this group, but low emission, excessive grid current and low-leakage resistance are also present. These failures are considered due to wear-out or old age, and should be predictable, since they are not the sudden-death type of failure. An adequate preventive maintenance plan would remove these tubes from service before an emergency break in operation occurs.

The last bar represents those tube returns which show no defect on our retest, that is, tubes still within manufacturing specification limits on the parameters tested. There are several reasons why tube returns may fall in this category: (1) Circuitry which—either due to design or faulty components—will not accept tubes that are still within limits; (2) use of a tube tester that does not test some parameter important to particular circuits; (3) precautionary removals by the maintenance man due to elusive troubles within a circuit or related circuits; (4) wholesale tube replacements during major overhaul or renovation of equipments; and (5) disagreement in findings between tube-tester equipments at the military base and those at our laboratory regarding tubes which were removed as falling below specified limits.

It is believed that at the base we are discussing, the fifth reason is the predominant factor in the "no-defect" returns. When a sample of the tubes in this category were retested on the basis of JAN specifications, they were found to cluster at, or very near, the lower transconductance limit. A distribution of the measured G_m values for type 12AT7 tubes in the combined electrical and "no-defect" categories is presented in Fig. 2. Only the G_m value of the low section is shown. A similar distribution is shown for the 6SK7 in Fig. 3. Plots of other tube types indicate similar results. The distributions in Figs. 2 and 3 indicate that the tubes in the electrical and "no-defect" categories are actually the same type of removal and that any cure for the electrical type failures would automatically eliminate most of the removals in the other category.

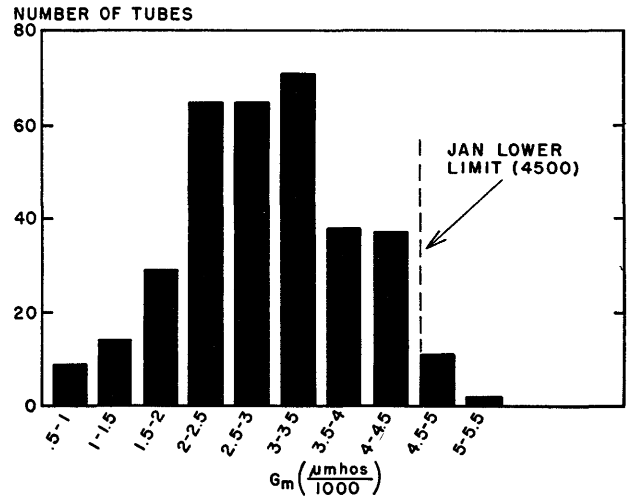


Fig. 2—G_m distribution of 12AT7 removals from land-based teletype equipment.

If the returns are grouped into the two general categories of mechanical failures and electrical deterioration failures, the percentages for this military base are 9 per cent mechanical and 91 per cent electrical. Improvement programs now in progress on military tubes would reduce the number of mechanical failures. Data from one Navy base show a reduction in this type of failure greater than four-to-one for one-tube type, as a result of the use of improved types.

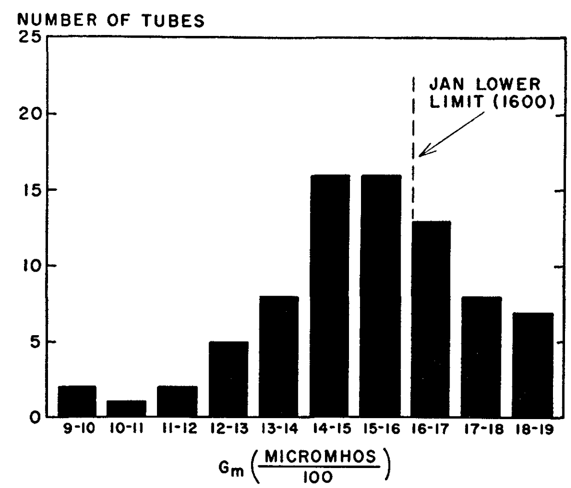


Fig. 3—G_m distribution of 6SK7 removals from land-based communications receivers.

More interesting than the mechanical failures are the 91 per cent of the tubes which fall in the electrical defect category. In the present state of the art, it seems safe to say that tubes in operation will eventually fail due to cathode-deterioration, deposits on the micas, cathode-interface formation, or other changes within the tube.

It should be mentioned that tube manufacturers are devoting much effort to the task of providing longer-life tubes. Some improved types incorporate slotted micas, coated stems and more passive cathodes. At this time we do not have life data on these improved types.

If we consider the loss in G_m that would result from a gradual degradation of the cathode emitting material, the results might be something like the theoretical plot shown in Fig. 4. Here we have assumed a normal distribution around the center value of JAN specification limits as the starting point of a lot of new tubes. This would vary from lot to lot. It has also been assumed that the G_m distribution will drift downward with time, and the spread between upper and lower values will increase. If a continuous check were made on these tubes in operation and if tubes were removed as soon as the G_m fell below the lower specification limit, a plot of the frequency of tube removal would be similar to the curve in lower part of figure. This is a skewed distribution, since the original assumption was that the G_m slope was accompanied by a spread in upper and lower values.

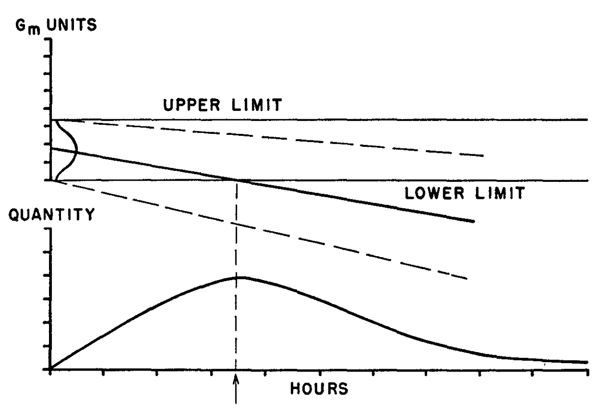


Fig. 4—Theoretical plot of expected tube removals based on G_m deterioration with time.

For individual tube types, the slope of the G_m drift would be different, but, once established, would allow a fairly valid prediction of tube usage rates in environments where tubes are allowed to wear out normally.

Fig. 5 is a plot of the average G_m decay of a group of tubes from our life-test rack. It is apparent that there is an almost linear decay with time, out to 6,000 hours. The lower line on the graph represents similar data supplied by a manufacturer on another tube type. In neither case is there any evidence of increased rate of fall during the period of the test.

How well the actual tube removal rates fit this theoretical curve is indicated in Fig. 6, which shows the quantity of 6BA6 tubes removed in consecutive 500-hour periods out to 10,000 hours. At first glance, the distribution does not appear to fit the expected curve very well. For example, there are too many returns in the first 2,000 hours. This suggests that the tubes should be classified into two groups on the basis of cause of failure. We do not know at this time the basic reasons for this group of early failures.

Another deviation from the expected curve is in the period between 2,000 and 4,000 hours. Here the number of returns is low. It should be remembered, however, that the tubes are not under continuous test and many

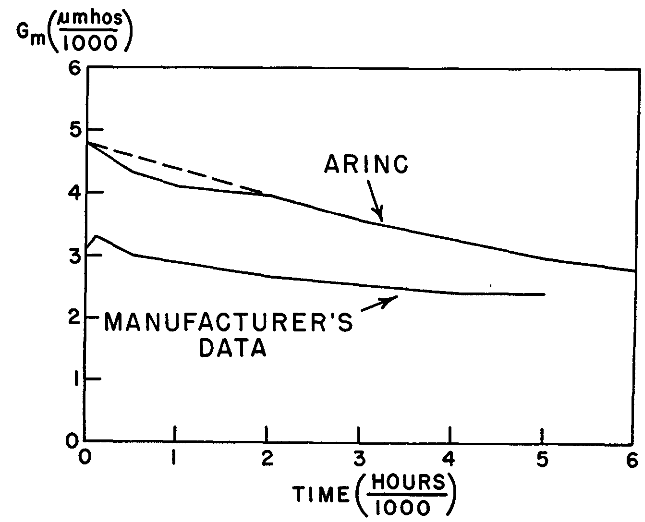


Fig. 5—Average G_m decay with time, based on laboratory life tests of two tube types.

of them drop considerably below the lower limit before being removed. This is illustrated in Fig. 7, which shows the actual G_m values of removed tubes vs. hours to removal. It will be noted that considerable spread exists in the G_m values at the time the periodic maintenance check is made. If the tubes with low G_m had been removed at the time they actually dropped below limits,

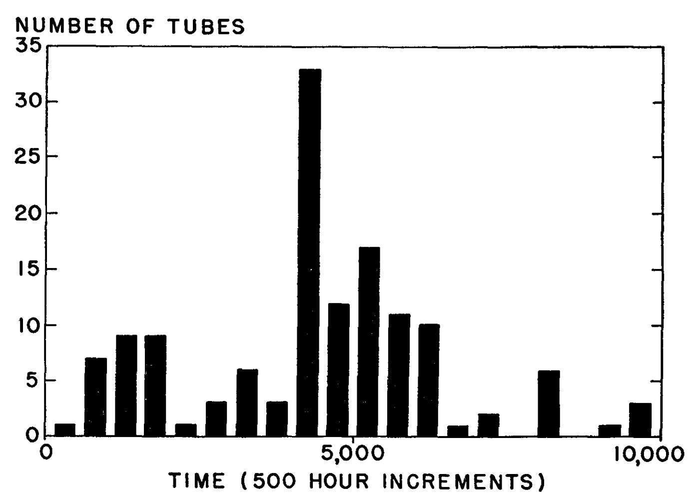


Fig. 6—6BA6 removals from land-based receivers by 500-hour periods.

the histogram in Fig. 6 would more nearly follow the theoretical curve. An estimated slope for G_m for the 6BA6 in this particular environment and application would be 450 micromhos per 1,000 hours, or a 10 per cent drop in the first 1,000 hours.

A histogram for the 6SK7 in similar use is shown in Fig. 8. Here, as before, the peak in removals is too pronounced, and the quantities of removals prior to that time are low. Again, these deviations are due to the interval of time between tests on the tubes. For the 6SK7, an estimated slope of G_m vs. time would be 70 micromhos per 1,000 hours. This amounts to a 3.5 per cent drop in the first 1,000 hours.

It should be pointed out that only with tubes of a

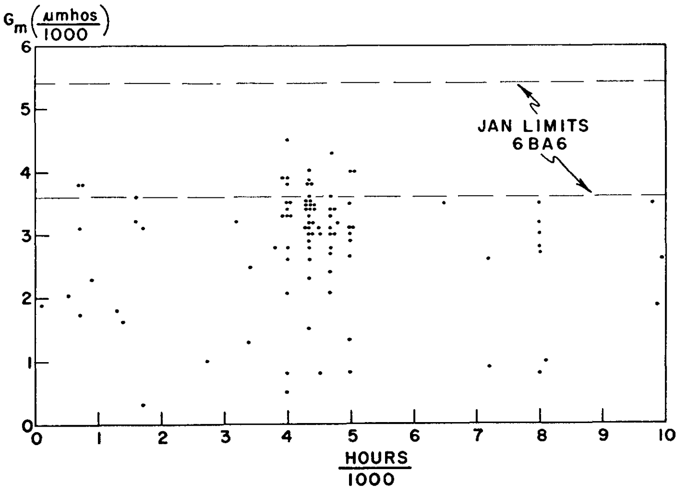


Fig. 7—6BA6 removals from land-based receivers: G_m vs hours of operation.

single type used in similar applications would this rise in failures show up at a particular time. Should many tube types or many applications of a single tube type be considered, there would be different slopes for G_m in each case, and the combined effect would be to level out the sharp rise seen in Fig. 8.

Preliminary data on the 12AT7 tube indicate a trans-

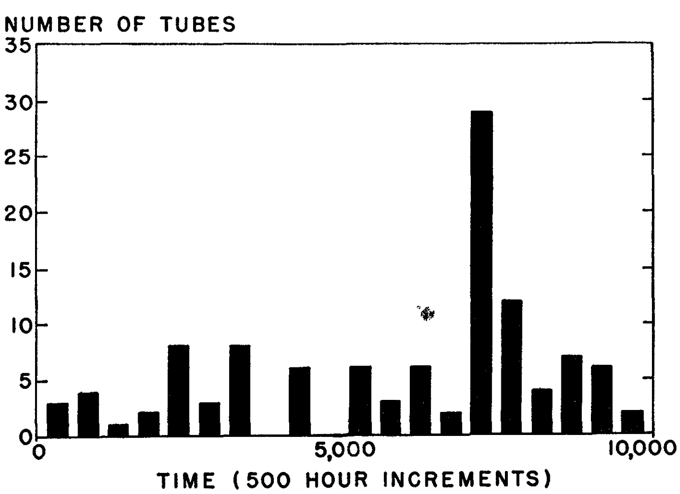


Fig. 8—6SK7 removals from land-based receivers by 500-hour periods.

conductance drop in excess of 700 micromhos per 1,000 hours. For this tube type, the pattern of removals versus time is complicated by the fact that malfunction of only one side of the tube is enough to cause removal of the entire tube. The effect is to remove a good section from our test because the other section is bad. This spreads out the expected increase in removals indicated in the theoretical curve shown in Fig. 4. The expected average time to removal is therefore more difficult to predict.

A controlled test has been installed in teletype terminal equipment to compare performance of the 12AT7 and an improved version of this type, the 6201. Since a majority of the tubes are still in service, final conclusions cannot be drawn. The test has been in operation 4,400 hours, and out of 33 tubes of the 12AT7 type installed, 11 have been returned as failures, against a return of 4 out of 36 for the 6201. Very limited data indicate the decay of G_m with time in the 6201 to be about one-half as great as the 700 micromhos per 1,000 hours observed in the case of the 12AT7.

We also have in operation in the same equipment 31 tubes of the 6SN7GT type, for the purpose of studying the build-up of cathode interface resistance. This group of tubes contains about equal quantities of cathodes of three types, that is, active, normal, and passive. Previous returns indicate that 95 per cent of all 6SN7GT tubes had measurable interface resistance. After 4,400 hours, none of the controlled tubes have failed; therefore, we do not yet have information indicating which type of cathode will give the best performance.

Cornell University measurements on a sample of 942 tubes of 14 types from the base under discussion show that 82 per cent have measurable interface resistance. Considering that 9 per cent of the tubes from this base have mechanical defects, it would appear that approximately 91 per cent of the electrical-type failures have some interface resistance.

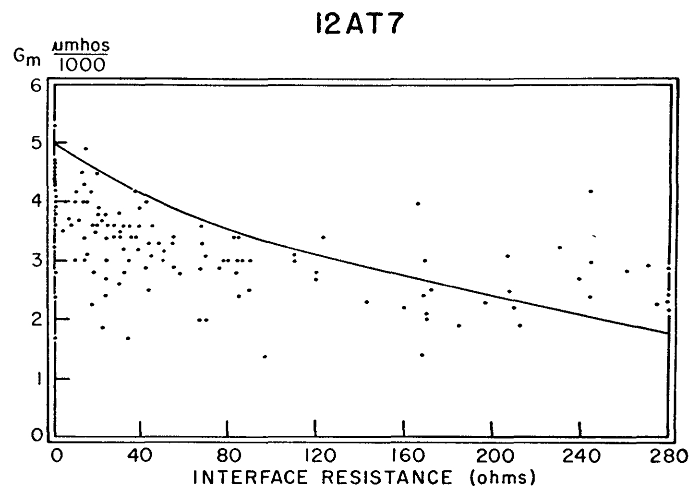


Fig. 9—12AT7 removals from land-based teletype equipment: cathode interface resistance vs G_m.

Fig. 9 represents a plot of measured interface resistance vs the measured G_m of some 12AT7 tube returns. Each dot represents the two measurements on one section of a 12AT7. The solid curve is a calculation of G_m vs unbypassed resistance in the cathode of a tube. For this calculation, an original G_m was chosen from the center value of a group of new 12AT7's. New tubes tested were not distributed around bogie, and a lower value was used for this calculation. The tube returns should have been distributed on both sides of the solid curve if the only cause of low transconductance was the effect of unbypassed resistance in the cathode circuit. Since the readings on the tubes place most of them below the line, it is assumed that some other deterioration of the cathode has also occurred. The nature of the resistance is such that it varies with cathode current; thus, the measured value may decrease considerably during test if appreciable cathode current is used. This accounts for the large values of interface resistance plotted above the calculated line.

Shipborne Equipments

The second group of tube returns which I shall discuss are from shipborne equipments and, thus, from a different type of environment than those previously considered in this paper. As would be expected, the characteristics of these returns are also different.

The equipments in this case represent many types of application, such as communications, radar, sonar, and others. For the most part, they are rack- or console-mounted, and subject to only limited vibration and shock. Shock mounts are provided where shock is thought to be a factor. A system of periodic preventive maintenance is in effect, although the frequency of this maintenance varies from equipment to equipment and from ship to ship. In general, daily operational checks are made on equipments which are in more or less constant use—that is, 500 to 700 hours per month—and monthly tube testing is performed when practical. Wholesale tube testing usually occurs prior to the start of an extended cruise.

It would seem that tube testing plays a more important role in the maintenance of complex equipments than of relatively simple units, such as audio amplifiers and communication receivers. Where troubles tend to be elusive, as in the case of complex equipment, greater significance is attached to the minimum readings indicated on tube testers. This may result in frequent tube testing and, consequently, more handling.

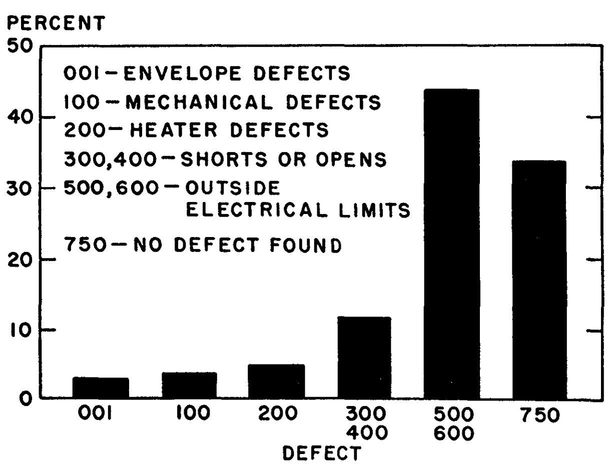


Fig. 10—Defect distribution of tubes removed from shipborne equipments.

The defect distribution of tube returns from the shipborne equipments is shown in Fig. 10. Comparison with the similar distribution for land-based equipments (Fig. 1) shows an increase in all types of mechanical failures. Whereas 9 per cent of the returns distributed in Fig. 1 were mechanical-type failures, Fig. 10 shows a total of 22 per cent in this classification. The greatest difference is in the fourth bar from the left, representing shorted or open tubes, which in this case accounts for 12 per cent of all returns, compared to 2 per cent in Fig. 1. This difference may be due to factors other than shock and vibration during operation. For example, where there is insufficient storage space for new tube stocks aboard ship, the tubes in some cases are removed from their cartons and stored in sliding drawers. Increased handling or testing for shorts by standard procedure of rapping the bulb probably increases number of tube returns.

The fifth bar in Fig. 10, representing the tubes testing out of electrical limits, is considerably lower than in the distribution for the land-based equipments. In this chart, 44 per cent of the tubes fall in this category, compared to 63 per cent in Fig. 1. Each of these percentages, of course, is relative to the other defect percentages for that particular base. A reduction in the electrical failure group usually indicates that tube removals, on the average, have less operating time. Application, environment, and maintenance procedures combine to reject the tubes for reasons other than normal wear-out. This condition is also reflected in an increase in the last bar on the chart, which represents the tubes in which we found no defect. The increase is from 25 per cent in Fig. 1 to 34 per cent on this chart.

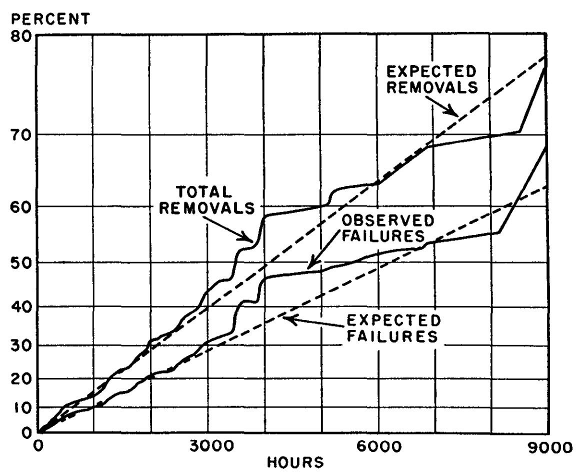


Fig. 11—Tube mortality with time, shipborne equipments.

A controlled test was initiated in shipborne equipments, using 44 tube types in all applications available. The controlled tubes installed represented a cross section of the JAN tubes currently used. Fig. 11 shows the mortality rate of all tubes in this test. The upper curve represents the total removals at each time interval, expressed as a percentage of the total number of tubes in the test, and the lower curve represents the portion of the total returns which exhibited some defect on retest. The reason for the lower curve is that "no-defect" tubes from this base cannot be considered in the same group as the electrical failures, as was true of the returns from the other base discussed. Instead, these tubes which "test good" are scattered within the specification limits in about the same pattern as new tubes. As a further indication of their characteristics, a sample of 165 of the "no-defect" tubes were reinstalled in sockets identical or equivalent to those from which removed, and only 2 failed to operate satisfactorily. Some of the reinstalled tubes have now operated 3,000 hours without failing.

It is assumed that most of the removals of satisfactory tubes are the result of emergency maintenance practices under which more tubes may be changed than necessary to effect repair or the tubes are changed first and the real source of trouble traced to some other component afterward. Tubes removed under such circumstances are seldom reinstalled. In any case, it appears that tubes in this category did not cause an interruption of service and should not be considered in calculations of equipment reliability.

On both curves in Fig. 11, the dotted line represents a calculated exponential that most nearly fits the measured data. It is noted that the measured result follows the calculated curve very closely for the first 3,000 hours. The hump in the measured curve between 3,000 and 4,000 hours is traceable to a maintenance period during which a group of 6AG7 tubes found to be below specification limits were removed.

It is possible that the maintenance period in question corresponds to the point at which—due to the normal G_m drift for the 6AG7 type—the large quantity of tubes at the center of the distribution would cross the lower G_m limit line. It is surprising that this would be detected, considering that so many tube types were included in the test and that different types would have different G_m decay rates. In this test, relatively large quantities of 6AG7 tubes were installed, and these tubes determined the failure rate at this point. Had the tubes been installed at staggered time intervals, this rise in failures would have distributed along the curve more evenly, and the measured data would more nearly fit the calculated curve. The test was made on a lot of new tubes, and a removed tube was not replaced with another controlled tube. Thus the quantity of tubes in the test decreased with time. Actual equipment operation demands tube replacement with failures, and at any period an equipment would have random operating times accumulated on each tube. Under this condition, the expected failure rate might be higher than indicated here.

The conduct of this controlled test was hampered by termination of controlled tubes due to shifting and modification of equipments after the first few thousand hours. Beyond 7,000 hours, too few tubes were left in the test to provide reliable data. If we can say that the calculated exponential failure rate shown is a good first approximation, it is possible to make predictions on equipment reliability due to tubes. The slope of the failure rates could be established over relatively short periods of time.

Similar curves may be drawn for aircraft environments; however, hours of operation accumulate much more slowly in these applications, and we do not have data in the thousands of hours for aircraft equipments. In general, the returns up to 500 hours indicate that the percentage tube-reject rate is greater than in the environments discussed in this paper.

SEAC—Review of Three Years of Operation

P. D. SHUPE, JR.† AND R. A. KIRSCH†

Purpose for Construction of SEAC

SINCE SEPTEMBER of 1950, the National Bureau of Standards has had SEAC, a digital-automatic computer, in almost continuous daily usage. It was originally conceived as an interim-computing facility for the use of the government until a more complete computing system could replace it. Consequently, in constructing the machine, it was intended to put into productive operation as soon as possible a minimal machine that could produce computed results. However, the machine proved quite reliable, and the experiments involved in its design were sufficiently successful that SEAC was expanded and kept in operation as a permanent tool at the National Bureau of Standards. It is the purpose of this paper to present some of the operating experience that has been obtained from the use of this computer and to indicate the ways in which component reliability and maintenance procedures have affected the amount of useful computation that has been obtained from SEAC.

When SEAC was first put into productive operation, the demand for its use was so great that it became necessary to schedule it for operation on a 24-hour-a-day 7-day-a-week basis. Since it was apparent that the computer would have to be expanded, at first this time was divided nearly equally between the engineering groups modifying the computer and experimenting with it, and the mathematicians who were producing computed re-

† National Bureau of Standards, Washington, D. C.

sults for the various Government agencies. After a year of operation, the engineering time was reduced to about one-quarter of the total available time.

By the use of flexible scheduling, and largely because of the very convenient input-output facilities of SEAC, a great number of concurrent mathematical projects were able to use the computer. This kept a staff of from 30 to 40 mathematicians busy formulating and coding problems. It has at all times been possible to allow both short and long runs on the computer because the time that is involved in switching over from one problem to another can be less than 2 minutes. This has resulted in SEAC being used for over 85 different projects of varying length and of diverse natures.

Interspersed with the productive computation have been periods during which research and development have been conducted on SEAC. These engineering periods are, typically, eight hours in length. During this time, investigations have been conducted into new computer circuitry and accessory devices. As a result of this engineering work, many new features were added to the computer.

Because SEAC has been used for both computation and development, there has been some decrease in the reliability of the system from what it would have been on a SEAC that was solely a computing device. Examination of the operating record shows that since the initial period of experimentation, this loss of reliability has been small. After engineering periods, the attempt is made to restore the machine to the condition it was in prior to the period. Obviously, this is not always possible. In general, however, the engineering periods on the machine do have the effect of hastening failures that may be intermittent or marginal. Removing components, turning power on and off frequently, and physically dislocating sections of the circuitry for the duration of temporary experimental changes accelerate the failure of components that would perform satisfactorily for a longer time in ordinary operation. An analysis of machine failures, which will be given later, will show the extent to which experimentation has caused failures in machine operation.

Operating and Maintenance Procedures on SEAC

Until recently, when the 168-hour-a-week schedule was relaxed, there were four periods per week of approximately eight hours in length during which SEAC was used for engineering or during which maintenance work was performed. At all other times, the computer was operated by mathematicians. These mathematicians are responsible for the mathematical formulation of the problem, for the preparation of the code, and for operating the console during the running of the problem on the computer. In general, they are not expected to be familiar with the logical structure of the computer except insofar as it is manifest in the structure of the operation code.

When a problem is scheduled for solution on SEAC, the mathematician who has prepared it puts the problem into the computer and follows its operation during the time that it is on the machine. Since there is no automatic-checking provision in SEAC, with the exception of the memory-parity check, it is usual for the mathematician to provide program checks or to have printouts on the teletype at sufficiently frequent intervals to enable him to monitor the operation and to detect any machine errors. It is also usual when the mathematician suspects a failure in machine operation for him to call a maintenance technician. The information that the mathematician can provide is usually sparse if not misleading. Occasionally, the mathematician gives a report simply that the machine has "hung up." The technician must then determine whether the machine has indeed made an error or whether the error is in the code. Not being familiar with the code, he often first tries to rerun the section that is claimed to have produced the error. If inconsistent results are obtained, it is fairly certain that the machine has made at least one intermittent error.

The maintenance staff of SEAC for three-shift operation consists of three engineers and five technicians. One technician is occupied full time in the construction and repair of replacement parts. At all times when the computer is in operation, there is a technician present and an engineer available for consulting purposes.

In the event of a machine error, it is desirable for the maintenance technician to be able to reproduce the machine conditions under which the error occurred. Unfortunately, the technician is generally not qualified to analyze the code that was running and to detect the immediate nature of the error. In such circumstances, it would be highly desirable if the person who operates the computer were familiar not only with the code but also with the logical organization of the computer. For a machine like SEAC, where the electronic structures are highly iterated, it is necessary only to have a machine operator who can analyze troubles from the logical standpoint and a technician familiar with the electronic nature of the circuitry in order to maintain the machine. The luxury of a standby engineering staff present for consultation in emergencies is a fortunate aspect in maintaining an experimental machine. For computer installations of a nonexperimental nature, this auxiliary staff is not available.

In the event of a failure, the computer is not always removed from problem solution. When the failure is highly intermittent, it is usually more efficient of time to allow computation to proceed until such a time as either computation becomes impossible, the error occurs frequently enough to make it possible to locate it, or a scheduled maintenance period occurs.

In the diagnosis of an error, several systematic procedures are used. The most frequently used technique is the diagnostic-test routine. There is a library of such routines available to the person doing the debugging.

In general, the test routines are predicated on the assumption that the operator has assured himself that control functions in the computer are operating properly. These routines then perform any one of a number of computations involving special parts of the computer with a diagnostic printout if an error is produced. One routine, for example, loads the acoustic memory with different word patterns and then checks the memory for the storage of these patterns. If an error is detected, the routine prints on the Teletype information that indicates to the technician operating the routine where the failure has occurred in the memory. Obviously, this indication can only be approximate, but for errors that involve such commonplace failures as those due to improper gain adjustments in the recirculation amplifiers, a test routine which will save diagnosis time for technicians is highly valuable.

There is a compensating disadvantage in unqualified use of diagnostic-test routines like this memory test. When technicians and semi-skilled maintenance personnel use these routines, they have a tendency to rely too greatly on the indications provided by the test routine. When subtle troubles occur for which the test routines were not designed, there is often a significant loss of time involved in trying to find failures where they do not exist. Despite the occasional lapses into indiscriminate use of these test routines, the amount of effort saved by allowing the computer to do its own testing is great.

Among the various diagnostic tests available for use with SEAC are those that check specific portions of the machine: the memory, either acoustic or electrostatic, the arithmetic unit, the magnetic-tape auxiliary memory, and the magnetic wire input-output. Other routines cause the computer to perform operations which result in the highly repetitive production of special patterns of standard pulses at test points. These routines are used in conjunction with an oscilloscope for observing the patterns produced. Failure is then detected visually.

Another type of test routines frequently used is written as the trouble is observed. It is usually less complicated than the diagnostic routine and is written to test for a very specific trouble. It is also more effective than the diagnostic routine for troubles that involve the control section of the computer and other troubles of a serious nature that cause radically incorrect behavior of the machine. This type of routine is also of use in the detection of highly intermittent errors where the lower-duty cycle of testing of some diagnostic routines might make the detection of the error less probable.

Another systematic procedure for the detection of computer errors, which has always been in use on SEAC, involves marginal checking. In SEAC, most of the signal outputs of tubes are coupled to the rest of the circuitry by the use of pulse transformers. By varying the dc voltage to which these transformer secondaries are returned, it is possible to vary the effective output voltage of all tube and transformer stages in the computer. Only two such voltages need to be varied to affect almost all stages in the computer in the same manner. These two voltages can be used to provide an over-all marginal test of the computer or of individual chassis. It is often possible to set these voltages at such a point that only the single weakest stage in the computer will be effectively inoperative. It is also sometimes possible to increase the frequency of intermittent failures by this technique. The marginal check is incorporated as a part of the preventive-maintenance schedule.

With one single area of the computer failing under marginal-voltage variations, the trouble is traced to the individual component that is to blame. Note that all during this period the operation of the system would be error-free under normal voltages. This forcing of the system to fail allows trouble shooting to be performed on the computer during "cheap" time, that is, during a maintenance period rather than during time that would otherwise be scheduled for computation. For components that are approaching the failure condition gradually, as in the case of vacuum tubes whose transconductance may gradually decrease, it is possible to anticipate failures during maintenance by marginal checking. The marginal check gives a nondestructive quantitative measure of the operating tolerances under which the machine is working. Another reason for the marginal check is to allow computer troubles to be debugged one at a time. If the failures are allowed to accumulate between maintenance periods, the situation will arise where there are two or more faults present in the computer simultaneously. The difficulty in locating the source of malfunctioning under such conditions is vastly greater than the effort that would be needed to isolate them individually. The computer represents a very powerful tool for use in debugging many parts of its own internal structure. However, to allow the computer to lapse into the degree of disrepair in which more than a single trouble is present at a time is to make it generally very difficult to use this powerful debugging tool.

For over two years of operation, SEAC had no automatic-checking facilities. All checks that were performed were programmed. For example, one checking procedure that was devised enables the computer operator to minimize the amount of time that is lost in the event of the detection of an error. Once every half-hour or hour during long runs, the entire contents of the high-speed internal memory are recorded on a magnetic-wire unit. Under the control of the program, this recording is then read by the computer and a check sum of the recording is compared with the corresponding check sum of the contents of the memory. If the sums agree, it indicates that the recording has been made accurately, and the machine automatically resumes computation at the point in the main routine where it left off. The memory-recording routine requires only eight memory cells and is completed within less than two minutes when transferring the contents of the 1,024-word memory. If the operator makes use of this routine, he can insure that in the event an error is detected, the time lost will be no more than the time since the last memory recording was made. To resume computation at the point of the last memory recording, he simply reads the recording back into the memory. In a matter of a few seconds, the computer is recalculating in the part of the routine that occurred after the last recording. Not only is this routine useful for minimizing lost time due to errors, but it enables short periods of machine time to be used effectively in the solution of long problems.

At the beginning of this year it was decided to incorporate some degree of automatic checking in one section of the computer that gave less reliable operation than other parts of the electronic circuitry. It was found that transient errors involving the change of a single binary digit would occur in the acoustic memory. Often, these errors could not be attributed to the failure of any single component. On the other hand, such errors were relatively rare in the main body of the computer which used more conventional circuitry. This was a case in which checking circuitry could be built that would have a margin of reliability considerably greater than that of the circuits being checked. The parity checker that was built incorporated standard SEAC-type tube and transformer stages with diode gating. After its initial experimental stage, this checking circuitry was able to detect the great majority of errors in the acoustic memory. Because of the experience and success gained with this addition to the computer, investigations of the possibility of incorporating automatic checking for the input-output and electrostatic-memory circuits are now under way.

Analysis of the SEAC Log

Whenever there is a failure in the operation of SEAC an entry is made in the *Operations Log*. These entries may be made either by the operator or by a technician or engineer. A record is also kept of all modifications to the computer. Because of the transient nature of some errors that the computer makes, it is not always possible to identify the cause with certainty. Therefore, it is often the case that an entry will simply record the loss of computing time with no explanation of the cause. At the end of each week, figures are obtained for the operating efficiency of the computer. Operating efficiency is defined as the ratio of productive computation during assigned time to total assigned time. The remainder of assigned time after productive computation constitutes machine errors, overrun of engineering time into scheduled operating time, and downtime due to debugging.

A graph of the operating efficiency of SEAC for three years of operation is shown in Fig. 1. The average efficiency for that time was 74 per cent. The ratio of code-checking time to productive computation time is also shown. Only a small portion of operating time is needed for code checking because of such features as the automatic monitor and the high-speed wire output which enable each operation to be monitored as it is performed and to be recorded rapidly on an output unit for later transcription to printed copy with auxiliary equipment.

In order to understand the manner in which various failures have contributed to downtime on SEAC, the operation log may be analyzed. Naturally, some of the failures that are of a transient nature cannot readily be

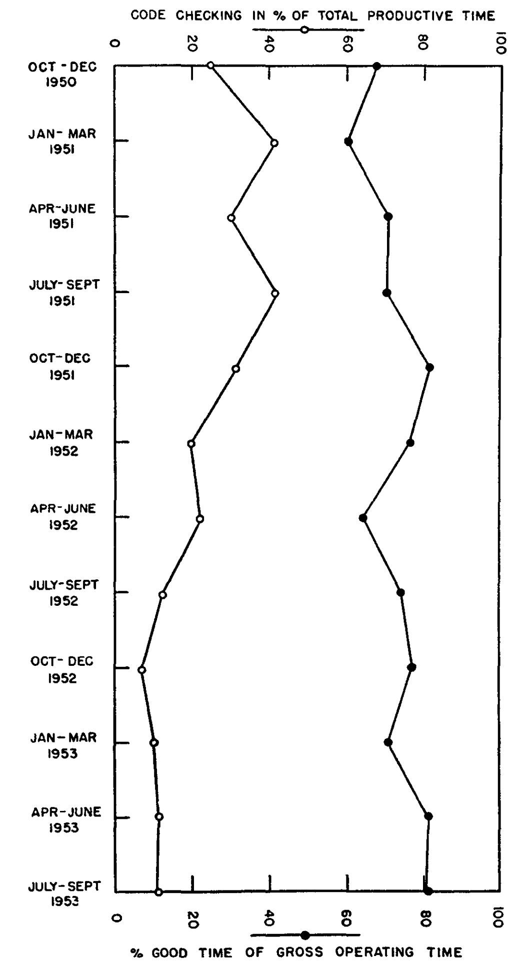


Fig. 1—Graph of operating efficiency of SEAC from October, 1950 to March, 1953.

analyzed. However, a recent period of a month has been selected during which a large percentage of the failures that occurred during scheduled operation were capable of analysis. Of a total of 477 hours scheduled for computation during that month all but 112 hours produced error-free calculation. The operating efficiency of this period was therefore 77 per cent. Only about 12 hours of lost time could not be attributed to specific faults. Data from this analysis of the operation log are in Table I.

In the first column is the cause of the failure. The second column gives the number of individual cases during which computation was delayed due to a component failure. The numbers in this column are considerably greater than the total number of times that debugging was necessary because in many cases the machine errors were of a trivial nature that were easily detected or corrected by the operator. The third column gives the amount of downtime that was caused by these machine errors. The fourth column of the table gives an indication of the most common length of computation time lost as a result of the failure. The last column gives the number of individual components the failures of which caused the number of computer failures shown in the second column.

TABLE I

NATURE OF FAILURES ON SEAC FOR ONE MONTH

Cause of Failure	Number of Machine Failures	Total Downtime Hrs.-Min.	Most Common Length of Downtime Hrs.-Min.	Number of Components that Failed
Diodes	8	6:05	0:30	3
Vacuum tubes				
6AN5	3	2:30	0:45	2
6AK5	3	2:35	0:45	2
Input-Output and Magnetic Tape	29	14:20	0:30	—
Pulse Transformers	14	17:30	1:00	3
Experimentation				
Parity Checker	26	11:40	0:25	—
Other	10	9:05	0:55	—
Electrical Connections	4	6:50	1:30	4
Switches	2	3:30	1:45	2
Adjustments				
Acoustic Memory	7	2:10	0:20	—
Magnetic Input-Output Units	9	5:25	0:30	—
Mechanical Equipment	13	4:45	0:20	—
Auxiliary Equipment	18	5:20	0:15	—
Miscellaneous	5	6:40	—	—
Undiagnosed	20	13:30	—	—

The table shows that during this time there were three diode failures. These were bad diodes that were not found during the preventive maintenance diode check. There were eight individual times during which these three diodes caused machine failures. All other failures among over 15,000 diodes were detected during the preventive check before they could cause machine failures.

During this period of a month there were four tube failures that caused six failures in machine operation. The two 6AN5 tubes came from standard SEAC pulse-repeater stages. They were rejected for low emission. The two 6AK5 tubes were removed from the acoustic memory.

The most frequently occurring fault in SEAC operation involved the input-output equipment. In general, no component was at fault. Rather, mechanical variations that are difficult to control in such equipment were at fault. Because there is no automatic checking for the input-output, incorrect-data input often is not even detected after the computation has progressed to the point where programmed checks occur, although for ordinary reading of program information, a checking routine is common. This minimizes the lost time from such input. The most costly input-output failures occur when the final computed results are incorrectly transferred to the output medium.

The failure rate shown for transformers is not typical in SEAC operation. It does, however, serve to indicate the extent to which certain components cause loss of computing time all out of proportion to their frequency of occurrence. The transformers that failed during this time were all of the standard type used in the SEAC pulse-repeater stage. For over two years of operation, transformer failures were so rare as to make it unnecessary to do any checking of them. Furthermore, it was correctly anticipated that the most common type of failure that would occur was catastrophic in nature, and it is not easy to anticipate such failures by testing the suspected transformers. As a result, when these three transformers developed intermittent openings or shorts in their windings, they caused a great deal of trouble before they were located.

It has already been mentioned that some failures in SEAC could be attributed to its use as an experimental machine. Shortly before the month in consideration, it was decided to install a parity checker for the acoustic memory. The new circuitry was debugged and installed in the computer. It began to detect many "errors," some spurious and others of such a nature that they would not cause computer malfunctioning. Eventually an error was found in the construction of the new unit, but not before it had caused a great deal of lost time on the computer. Now that this circuitry has been thoroughly debugged, it is performing with the reliability that was anticipated and has succeeded in locating most of the transient errors in the acoustic memory. The remainder of the downtime attributed to engineering was caused by overruns of engineering time into scheduled computation and by changes that were made during engineering that were mistakenly not restored until after the engineering period had ended.

SEAC went into operation with a number of joints unsoldered. At first, there were few malfunctionings due to these poor connections, largely because the wires involved were uncorroded. As the wires became corroded, they caused malfunctionings and were consequently located. It is safe to estimate that there are still a few of these unsoldered or rosined connections in the computer, although they are probably in uncritical areas. Most of the current troubles with connections arise from those that have become loose because of the extensive removal of components. When plug-in components are removed, there is motion of the wiring on the plugs. In addition, for many of the voltage busses, there was inadequate provision to leave slack or stress loops in the wiring so that these connections have been rather prone to coming loose. During the month under consideration, there were four bad connections discovered, a fairly typical number.

In two cases during the month faulty switches caused machine malfunctioning. In terms of lost time, these were rather costly failures because the connections involved were intermittent. Most of the switches in SEAC are rotary type and cause little failure. The push-button and toggle switches have been less reliable but only because they are used more frequently and enthusiastically.

Although the next three adjustments shown in the table are performed during the maintenance periods, it has also been necessary to perform them between maintenance periods. The failures indicated generally occurred in groups in short periods before it was recognized that an adjustment was necessary, so that only about two of each type were necessary during the month.

A coincidental situation accounts for the large number of failures due to auxiliary equipment during this period. Magnetic recordings of codes and input data are generally made at times other than when the operator is scheduled to operate the computer. In this case many such recordings had been made but were not used on the computer until some time later. When the faults in the auxiliary equipment were discovered, it was too late to re-record the data and lost time resulted. As a result of this experience, it was suggested to the operators that they use another auxiliary device, the outscriber, to check all inscriptions on wire by punching paper tapes on the outscriber and comparing the output tape with the originally-punched tape. This is recommended where there are no programmed means provided for checking the input from magnetic wire.

During this month there were also five miscellaneous failures and 20 cases of machine failure not diagnosed.

Component Reliability in SEAC

Vacuum Tubes

There are a total of 1,424 vacuum tubes in SEAC and its associated auxiliary equipment, comprising 32 different types. However, since the type 6AN5 vacuum tube occurs more frequently than other types, discussion will be confined to this tube. A total of 1,050 6AN5 tube locations presently exist in SEAC. During the first three years of SEAC operation, approximately 2,500 6AN5's were used in the machine. Of these, 1,300 tubes were rejected for various reasons. Rejections were made almost exclusively during the preventive maintenance periods. Operational failures of 6AN5's in SEAC have been very few. During a 15-month period from February, 1952 to April, 1953, for example, it was necessary to replace only 18 tubes during computation time.

Fig. 2 shows graphically vacuum-tube survival for 1,775 6AN5's used up to March, 1953. This group does not include the approximately 700 tubes associated with the Williams memory and short experimental developments. The curve has been plotted by considering batches of tubes installed within 500 hours of the indicated average as single entities and weighting the points on a survival curve for such a group of tubes according to the number contained in each batch. This curve shows the percentage of tubes one would expect to survive after a given number of hours.

When SEAC was first placed in operation there existed considerable variation in heater voltage in various parts of the machine. Accordingly, plate currents were measured with heater voltages at 5.7 as well as 6.3 volts to allow for abnormal heater voltage during use in the computer. If there was a drastic change in plate current when the heater voltage was decreased, the tube was discarded as "heater sensitive." In addition to preventing weak tubes from being installed in stages having low heater voltage, it was thought that heater sensitivity might provide an indication that the tube would soon be rejected for low emission. Thus, in September, 1951, heater sensitivity was formalized in a specification which called for rejection of all tubes that showed a reduction of plate current of 25 per cent or more when the heater voltage was changed from 6.3 to 5.7 volts. This part of the specification was adopted even though at that time abnormal heater voltages had been corrected on all chassis. Approximately 55 per cent of all tube rejects in SEAC until June, 1953, were made for heater sensitivity. An analysis of tube data showed that after 8,000 hours of service heater sensitivity was the main cause for replacement. (More detailed information on tubes and diodes used in SEAC is included in a Special NBS Computer Circular, which is in process of publication.)

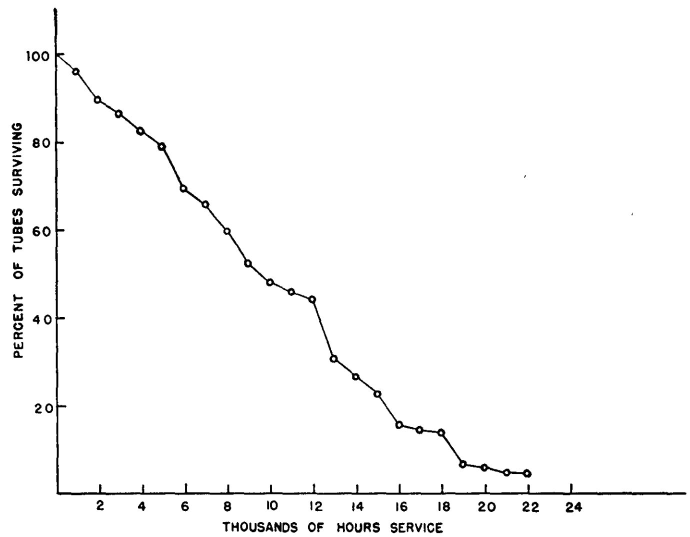


Fig. 2—Over-all tube survival versus time in service for 1,775 type 6AN5 tubes.

During June of this year, an analysis of our 6AN5 vacuum-tube experience showed that heater sensitivity increases with tube age and is some cause for alarm if heater voltages fluctuate. It is not apparent, however, that heater sensitivity provides a definite indication before a serious slump in plate current. The analysis also indicated that the median life expectancy for tubes rejected for all reasons except heater sensitivity was 10,000 to 12,000 hours, while if heater sensitivity was also included as a reject criterion, the median life appeared to be 8,700 hours. Since June of this year, the heater-sensitivity test has not been included in the vacuum-tube test for the computer. Experience does not yet show that this increase in tube-life expectancy as a result of relaxing the heater-sensitivity requirement coincides with any material increases in the incidence of tube failures during scheduled machine operation.

Germanium Diodes

The basic measure of diode reliability in SEAC is the rate at which diodes cause machine failures. Since there are over 15,000 diodes in the computer, any substantial failure rate would be intolerable. During the three years that SEAC has been in operation, however, this failure rate has been kept so low that diodes have been a minor consideration in the production of machine failures. Over the three-year period, diodes have had an in-operation failure rate of less than two per month. Table II shows the type of diode failures that occurred during a two-year period of regular machine operation. Those diodes with high forward voltage had a voltage drop of greater than 2 volts when a current of 20 milliamperes was passed through them. High back current was indicated by more than 500 microamperes when 40 volts was applied in the reverse direction. Drift was indicated by a change of more than 300 microamperes in reverse current while the diodes were under test.

TABLE II

OPERATIONAL FAILURES OF GERMANIUM DIODES IN SEAC
December 1950–December 1952
Total Diode Populations 15,676
(Elapsed Time: 3,693 to 19,512 service hours)

Nature of Failure	Number of Diodes
High E_f	4
High I_b	14
Drift	8
Unspecified	13
Total	39

The failures of Table II are few indeed and are offset by the preventive rejects shown in Table III. The figures for the computer exclusive of acoustic memory are representative of a period that is roughly half the period covered by Table II. The diodes came from sections of the computer (other than the acoustic memory) where environmental conditions were less harsh. The figures for the acoustic memory alone show the preventive replacements for a slightly shorter period than that of the computer alone.

TABLE III

PREVENTIVE MAINTENANCE REPLACEMENTS OF GERMANIUM DIODES IN SEAC
December 1951–December 1952
(Elapsed time: 10,830 to 18,905 service hours)

	Replaced for		Drift	Total Replacements	Total Population
	High E_f	High I_b			
Computer exclusive of acoustic memory	91	375	382	848	12,709
Acoustic memory	20	65	340	425	2,967

It can be seen from Table III that the diodes in the mercury memory suffered a much higher replacement rate than elsewhere in the computer. The memory data have been isolated in the tables because the conditions of inspection there were more rigorous and the conditions of operation involved higher operating temperatures than elsewhere in the machine. Inspections were made almost twice as frequently in the memory as elsewhere in the machine, thus slightly increasing the rejection rate.

The maintenance records of Table III do not include 276 replacements made in all parts of the machine during three short periods of poor operation in the summer of 1952. The air-conditioning equipment in those periods was inoperative during weather of unusually high humidity and temperature, operating conditions for which the machine was not designed. Ambient temperatures for the diodes ranged from 45 degrees C. to 55 degrees C. It was also necessary in this time to discard an unusually large number of "spare" diodes (used for replacements in the computer), probably because of the hot and humid weather conditions.

The weakest features of performance of the diodes are their tendency towards high back current and back-current drift. Actually, SEAC circuitry is extremely tolerant of back current; currents five times greater than specifications are allowable in most places.

It has been estimated that if diode specifications had been chosen to cover specific applications in the SEAC circuitry and grouped into three or even two types, the preventive maintenance replacements shown in Table III would have been materially reduced, perhaps as much as ten-to-one. Instead, at the time SEAC was designed it was decided in the interest of simplicity to use diodes to one specification—therefore all diodes were tested to the same specification as the ones used in the most critical part of the SEAC gating structure because the mounted configurations were not identified with circuit application. This tends to make diodes appear less reliable than they might actually be if typed according to application. This consideration becomes less important if preventive checking is not practiced and diodes are sought out and replaced only when failing.

Miscellaneous Components

In addition to tubes and diodes, four other types of components are used extensively in SEAC, namely, resistors, electromagnetic delay lines, capacitors, and pulse transformers. Of these, only resistors, delay lines, and pulse transformers form a part of the actual computer circuitry, capacitors being used solely to bypass places where excessive noise pickup would cause the generation of spurious signals.

Resistors in SEAC circuitry are required to be within 10 per cent of their design value. A set of measurements has been initiated to determine how far, in the three years of operation, resistors have varied from their initial rated value. The information obtained so far indicates that about 1 per cent of the resistors in the computer have exceeded the rated tolerance. These are being replaced as they are discovered. No records have been kept on the number of shorted or open resistors occurring during the three years. Two or three resistors are replaced each month during preventive maintenance.

Infrequent troubles with delay lines have been due to corrosion of the solder joint connecting the fine wire of the line to its termination. About a dozen instances of this kind have been reported during the life of the computer.

Pulse transformers have caused very little trouble in SEAC until quite recently when shorted and open primary windings increased the incidence of failure. Another type of failure was caused by decrease of transformer inductance which results from spreading of the gap after failure of the clamping band. Mechanical design of the transformers has been changed to correct these troubles.

Conclusion

The extent to which SEAC has served as an experimental machine during more than three years of operation is indicated by the increase in the number of vacuum tubes from about 750 to about 1,450. During a good part of this time SEAC was the only large-scale automatic digital computer available to the government. Its record of productive computation coupled with its expansion and increasing power and efficiency of operation have proved that it is possible to operate a computer for experimental purposes while obtaining useful computation from it.

The most recent of the major modifications to the computer has been the addition of automatic checking circuitry. The principle followed in the addition of this feature was that the reliability of the system should be balanced. Since the standard SEAC pulse-repeater stage has proved very reliable for continuous operation, it was the basis of construction of checking circuitry.

Maintenance procedures have had considerable effect on the reliability of the system. The large-scale removal of diodes and tubes has resulted in trouble from these sources being reduced to the point where they are among the least troublesome of the components in the computer. This has been accomplished at the expense of discarding some components that could have functioned in the computer for a longer time. If diode specifications were established on the basis of the circuit function for which use was intended, the number of rejects could probably be reduced by an order of magnitude without any loss in computer reliability. For vacuum tubes, the filament sensitivity test has proved to be too stringent, and by the elimination of this test, the number of tubes rejected can be decreased considerably.

Input-output has proved to be the least reliable of the functions that the system performs. However, recognition of this has greatly reduced this trouble by the incorporation of programmed checks into routines that involve large amounts of input-output.

New constructional techniques that have resulted from the experience with SEAC have been incorporated into DYSEAC. It is believed that the use of printed-circuit techniques will materially reduce the troubles that have resulted from the wiring methods that have been used in SEAC. In addition, the need for more rugged construction of removable components which has been indicated by SEAC experience has been incorporated into the DYSEAC design.

Three years of SEAC operation have helped make it a far more reliable machine than when first in operation. Weak elements have been eliminated, and the addition of new features have increased its power of operation.

Discussion

Harvey Rosenberg (Burroughs Adding Machine Corporation): How many control relays are used? What was their reliability?

Mr. Kirsch: In SEAC there are about 30 control relays. Most of them are associated with the input-output equipment. Most of the relays are used for gating low voltages—that is, the voltage at the level that comes out of our step-down transformers from the pulse repeater stage. So, problems like arcing and consequent corrosion of contacts are not very common with us. I do not have actual figures on relay performance but I should say that they are an extremely reliable device as we use them.

R. Kopp (Headquarters, United States Air Force): If, as I understood you to say, you try to continue production after an intermittent malfunction is detected but not cured, how do you know whether the resulting production is of any value?

Mr. Kirsch: Here we come to a problem with a nonoperational solution. It is very often a matter of what can, for want of something better, be called judgement. We cannot say ahead of time that we know for certain that the computer is going to be functioning well. However, if a mathematician suspects an error and if we respect the mathematician's suspicion, one thing to do is to put in a quick marginal check. This can be done quite rapidly on SEAC by the insertion of any one of several diagnostic test routines and the variation of the two voltages that I mentioned. If the failure is highly intermittent and is not one that is susceptible of increase by marginal variations, there is nothing that we can do but return the machine to the mathematician and wait until this error occurs again.

R. E. Lyons (Department of Defense): Has lowered filament voltage ever been used as a means of marginal checking SEAC? If so, was it an adverse effect on the 6AN5's?

Mr. Kirsch: Only recently did we introduce equipment into the SEAC for the variation of all the filament voltages, so we have not had any experience with that. However, in installing the 6AN5's in SEAC and in testing them during the preventive check, we used to give them a filament voltage test which consists of lowering the filament voltage from 6.3 volts, the nominal value, to 5.7 volts, at which value we measure the change in plate current. If there is more than a 25 per cent change in plate current, we say that the tube is heater sensitive. Our tube rejections over a period of about two years during which this test was used indicated that an increase in this heater sensitivity definitely occurs as the tube gets older. However, we have not been able to find any connection whatsoever between filament sensitivity and an imminent decrease in plate current, which is the important criterion in the SEAC dynamic gating stage. This heater sensitivity test has resulted in rejecting over half of the tubes that have been rejected preventively, but it has not proved to be a valid indication of imminent failure, so we have discontinued this heater sensitivity test. That is the full extent to which we have done any filament variations on the 6AN5's. Other tubes occur in such low numbers in SEAC that we have done no extensive experimentation on them at all.

A Review of ORDVAC Operating Experience

CHARLES R. WILLIAMS†

Introduction

The ORDVAC is one of three large-scale electronic computers located at the Computing Laboratory of the Ballistic Research Laboratories at Aberdeen Proving Ground, Maryland. It is the newest computer at the laboratory having been delivered in March, 1952. It operates in the binary number system in a parallel asynchronous manner, and it uses an electrostatic memory. Input to the machine is by punched teletype tape or punched IBM cards. Output from the machine is obtained on punched IBM cards, a teletype page printer, or punched teletype tape.

During the period from March 10, 1952 to October 2, 1953, the ORDVAC was available for computation a total of 8,410.18 hours. The total engineering time was 4,189.17 hours, and the remaining 1,104.65 hours was standby time. Standby time includes all the time during which no attempt is made to operate or service the machine. In one week the standby time is 168 hours minus the sum of the total engineering time and the total available time.

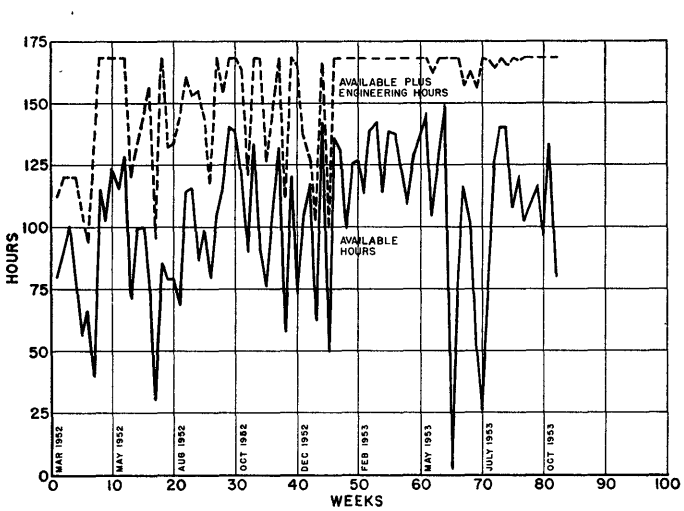


Fig. 1—ORDVAC weekly performance chart.

Machine Performance

The ORDVAC's performance to date is best described by the curves of Fig. 1. The lower curve is a weekly plot of the number of hours the ORDVAC had been available for immediate use. The period described is from March 10, 1952, when the ORDVAC was accepted by the laboratory, until October 1, 1953, a period of 82 weeks. The upper curve of Fig. 1 represents the number of hours each week that the machine had power applied to it. The distance between the upper curve and the 168-hour line thus represents standby time, and the number of hours between the curves represents engineering time.

During the earlier periods there was a considerable amount of standby time. The lack of sufficiently trained personnel and the lack of programs did not make it desirable to maintain the machine over the weekends. As more programs were completed, however, and the machine load increased, it was soon learned that operating the ORDVAC without interruption reduced the number of troubles encountered on Monday morning and enabled it to be made available for use at an earlier time. At present, the only time that ac power is removed from the computer is for repair of its cooling equipment, or because of a building power failure. Standby time has averaged 13.54 hours per week.

Fig. 2 illustrates the use that has been made of the ORDVAC by a comparison of the various classifications to which machine time is charged.

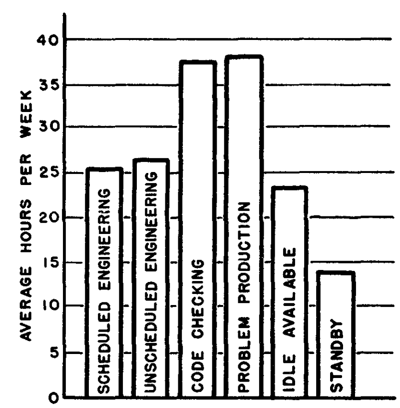


Fig. 2—Machine-time classification chart for an average week.

Within the period under discussion at present, an average of 37.17 hours per week was spent on code checking, 38.29 hours per week spent in production, and 23.03 hours per week classified as idle time. The title "idle time" is not intended to indicate that the machine is not in use during that time. The machine is made to run continually a test called Leap III, but it is available for other use at any time. Should the Leap III routine fail, the following time is immediately classified as unscheduled engineering until the cause of the error has been corrected, and the routine has successfully operated for fifteen minutes.

Recently, after all scheduled problems have been run, the remaining time has been used for problem duplication in an attempt to discover any errors that may have occurred during a regular production run. Normally

† Computing Laboratory, Aberdeen Proving Ground, Md.

problems are not duplicated unless suspected of being in error or, as just stated, unless time permits. Sufficient records have not been kept, but investigation has revealed that approximately 10 hours per week are lost as a result of errors that have occured. It should not be assumed, however, that the number of errors vary directly with the number of hours lost, because an error may not be detected until several minutes or even hours after it has occurred.

Perhaps one of the best yardsticks for measuring error frequency is the number of errors that occur over a given period while the machine is idle and running the Leap III test. During a three-month period in which the test was run 546.75 hours, the Leap III test failed 141 times due to the memory and 31 times due to the arithmetic unit. The average error-free running time was 3.2 hours. Again this figure is somewhat misleading in that the daily error frequency may vary widely. The time between the errors indicated above varied from a few minutes to a maximum of 24 hours.

Daily Schedule

The normal operating schedule includes three periods each day during which the performance of the ORDVAC is checked. The period between 0800 and 1000 hours is set aside for engineering purposes; at this time small changes and improvements may be made, troubles that may have developed during the night shifts are cleared, and, finally, the machine is tested by various self-checking programs. Tests are again made at the conclusion of each working day and at midnight. Upon the successful conclusion of the tests, the ORDVAC is released to the mathematicians who use it during the day. On the night shifts most of the production work is carried out by two technicians per shift who are in charge of the machine. They are not expected to understand the programs that are run, but they follow explicit instructions that are left by the mathematicians. This system has worked quite well. Usually, when a coder arrives at the laboratory in the morning he can expect to find that his program has been completed.

In the course of an average week 30 different problems are placed on the ORDVAC. Each problem may be placed on the machine several times each week. Ballistic computations, such as firing and bombing tables, and trajectories for guided missiles and rockets, form about 40 per cent of the work done by the ORDVAC. Vulnerability computations and data reductions form 25 and 20 per cent respectively. Research problems form 10 per cent and systems tests form 5 per cent of the work.

Testing the Machine

Three classes of tests are used to test the units of the ORDVAC. A read-around-ratio test is used for checking the memory; an input-output test is used to test the teletype and card-handling equipment; and the Leap III test already mentioned is used to test the arithmetic and control section and to give a further check on the memory.

Any electrostatic-type memory is subject to a fault which is known as read-around-ratio. In this paper, read-around-ratio refers to the maximum number of times that any one position in the memory may be consulted without causing errors in adjacent positions. A routine is used to scan the entire memory to discover any position or area that may be in an unnecessarily poor condition.

The present memory uses a three-dot system, which has the characteristic of being susceptible to failures of both ones changing to zeros and zeros changing to ones, should the read-around-ratio exceed a safe figure. Two tests are used to determine the read-around-ratio. They are similar except for the method of sensing a failure.

The routine is placed in the lower part of the memory from which it automatically proceeds to test the upper half. A number read into the machine from teletype tape determines the number of times each spot is bombarded before the adjacent spots (which had been previously cleared to ones or zeros, depending on the test) are scanned for failures. After half of the memory has been tested, the program is transferred to the tested area from which it proceeds to test the second half. Failures are indicated by a printed word identifying the cathode-ray tube involved and the address that was consulted. Up to a point, failures can usually be eliminated by the proper adjustment of the voltages on the cr tubes.

The input-output units are tested by a simple process of reading identical information into the machine from both teletype tape and cards and immediately punching out the same words. Comparison is made automatically, and discrepancies are printed out on the teletype page printer.

Perhaps the most thorough test is the Leap III test mentioned earlier. It is a revised version of the original Leap frog test written for the ORDVAC while it was still at the University of Illinois[1]; its name is derived from the manner in which it moves itself through the memory. In a period of about fifteen minutes, the routine moves in such a way that each order in it has occupied every memory address. Throughout the leaping process a self-checking system of arithmetic operations using pseudo-random numbers is carried out. If an arithmetic or storage error is detected, a set of twelve numbers are printed to reveal the nature of the error.

Occasionally the routine may fail in such a way that no useful information is obtained from the printed matter if, indeed, anything is printed at all. In this case a special routine is used to search the entire memory for the location of the program and then compare it with a correct copy.

It can be seen that the ORDVAC is used as much as possible to test itself by a highly repetitive use of its

[1] For a more detailed discussion of a similar test, see D. J. Wheeler and J. E. Robertson, "Diagnostic programs for the ILLIAC," Proc. I.R.E., vol. 41, pp. 1320–1325; October, 1953.

components. In this way, marginal and intermittent errors may be detected that would otherwise not be found. The usual memory errors are not to difficult to repair; the causes of arithmetic errors are somewhat harder to detect as their effects may propagate. Intermittent failures are especially troublesome if the error frequency is low. In this case an attempt is made to increase the error rate by using special routines which strain the circuits to their limits, by vibrating the circuit components, by varying filament or supply voltages, or by a combination of the above.

Engineering

It was mentioned before that during the nineteen month period under discussion the total amount of engineering work on the ORDVAC amounted to 51.36 hours per week. Broken down into scheduled and unscheduled engineering columns, the time was about evenly divided. They were 25.27 and 26.09 hours per week, respectively. Any engineering which necessarily interrupts the normal operating schedule without advance arrangement is classified as unscheduled engineering. In addition to the time taken each day to test the machine, scheduled engineering includes all the time which is taken for modifications to the machine. The major effort which has been made in the way of modifications to the ORDVAC has been devoted to the input-output system and to the memory.

Input-Output Modifications

The ORDVAC was delivered to the Ballistic Research Laboratories having as input a standard speed, five-hole teletype tape reader, and as output, a teletype page printer. Using this equipment, the time necessary to load the entire memory of 1,024 addresses was 38 minutes. The time necessary to print the contents of the entire memory was the same. No change has as yet been made in the page printer. It was realized quickly, however, that a change was necessary in the input in order to speed up the operation of the computer. Within a few weeks after the ORDVAC was placed in operation a control circuit had been designed and built for a modified tape reader which now allows information to be read into the machine at five times the previous rate; that is, the memory can now be filled in 7.5 minutes. This unit has performed exceptionally well, but input and output are still the big bottlenecks to more efficient use of the ORDVAC.

It is interesting to note the manner in which the University of Illinois has in the past used the ORDVAC remotely. The University would send to the Ballistic Research Laboratories a program through regular teletype channels which would be placed in the ORDVAC after insuring that the program had been received correctly. Answers were obtained on tape and returned to the University by mail or by the teletype channels. Thus the ORDVAC is available to any laboraory in the country having the necessary coding staff.

The most important change in input-output has been the addition of card-handling equipment. Before March, 1952, it was realized that card-handling equipment was desirable. The ORDVAC was to be placed in a laboratory with two other large-scale computers, the ENIAC and the EDVAC. The ENIAC already used punched cards, and such a system was being devised for the EDVAC.

There were two possible systems to be considered. The first was an external system in which information punched on cards was to be automatically converted to its binary equivalent by external apparatus. The second was an internal conversion system based on utilizing the ORDVAC itself to accomplish the necessary transformation. The second method was chosen.

The advantage of the internal system lay obviously in the elimination of the necessity for extensive additional equipment; and since established machine operations were utilized, the maximum in reliability was obtained. Another advantage was that less time was required to complete the addition of the equipment to the ORDVAC. The greatest disadvantage of the internal system was that a significant portion of the memory capacity was committed to the conversion program. The space required for both input and output conversions has amounted to 200 words, approximately one-fifth of the memory. Recent improvements of the memory, however, have made it possible to reduce the size of the program.

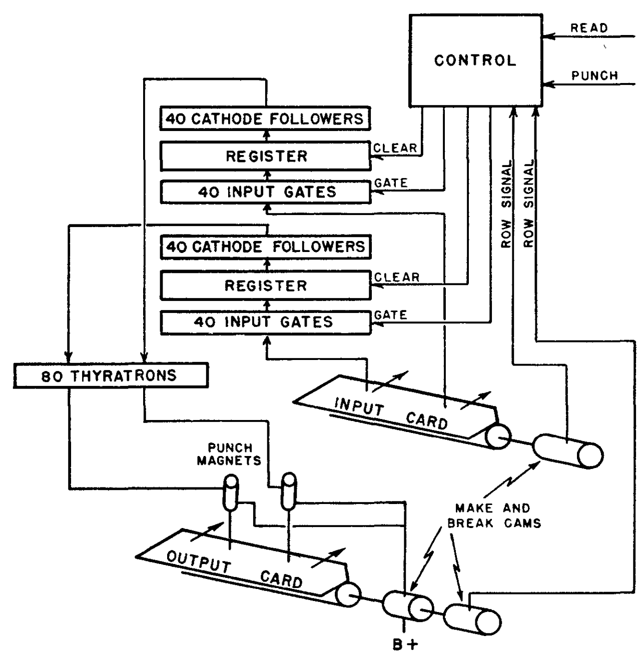


Fig. 3—Block diagram of the card input-output system.

The system is based on double-register gating in which the entire 80-column output can be gated to or from the ORDVAC simultaneously. A block diagram of the system is shown in Fig. 3. An electronic control circuit which is actuated by signals generated in the card punch or card reader allows the contents of the ORDVAC registers to be stored in the memory or changed in the interval between punching or reading successive rows on a card.

This system has proved quite reliable and is preferred to tape by the mathematicians. The usual practice is first to prepare a program on tape and immediately transcribe it to cards with a special routine that automatically reads from the tape and punches, in binary form, twenty-four words per card. Cards are used thereafter.

Several forms of input may be used. One form, just mentioned, in which 24 words per card are used, allows the memory to be filled in about 30 seconds. Another form of input, used for decimal data input and for which the 200-word input-output conversion routine is required, allows only eight decimal words to be punched on a card. Each card may be read into the ORDVAC in 1,033 milliseconds; this is equal to a rate of about 60 cards per minute. Data cards of this type may be used on either the ORDVAC or the ENIAC and may be handled with the usual card-handling equipment. By a third input method, 12 words are read from a card and stored in the machine by a repetition of an "order pair."

Memory

The electrostatic memory has always been the ORDVAC's weakest unit. Prior to June, 1953, the memory operated rather consistently at a read-around-ratio of 10 to 16. A restriction of 10 placed on the coders resulted in slower routines or routines that occupied a larger portion of the memory. In addition to read-around-ratio difficulties, there existed the problem of obtaining cr tubes whose screens were free of impurities. Type 3KP1 cr tubes were used for storage. Approximately 25 per cent of those tested were acceptable; most of the rejections were due to impurities which could cause improper storage. Until May, 1953, approximately 60 cr tubes were removed from the machine for various reasons: About two-thirds of the total number were removed because of the impurities in their screens; five were removed without proper cause; and the remaining tubes were removed because voltage adjustments could no longer hold the read-around-ratio of those tubes above 10. No cr tubes have burned out under normal operating conditions, although five were burned out when subjected to abnormal conditions.

In May, 1953, it was learned that the University of Illinois was obtaining improved read-around by using a three-dot system.[2] In this system, as used in the ORDVAC, two dots represent a one, and three dots represent a zero. The first and third dots occur at the same location.

The change over from a two-dot system was made about June 1, 1953. Considerable difficulty was encountered with the sensing pulse; and, in order to strengthen it sufficiently so that its excursion would not vary with the number of zeros sensed, it was necessary to incorporate a pulse transformer into the circuit. The new system caused the read-around-ratio to be increased by a factor of three so that it then varied between 32 and 48.

In the hope of further improving the reliability of the electrostatic memory, a number of type C73376B cr tubes, which are under development by RCA, were obtained and installed in the ORDVAC. At this writing they have been in the ORDVAC 1,500 hours and seem to be operating satisfactorily. The read-around-ratio is at present guaranteed to be 80 and is usually 100; the flaw problem no longer exists since these tubes are virtually flaw free. Although the reliability of the memory has been increased, the improvement is not very noticeable in Fig. 1. Unfortunately, a series of arithmetic errors, which were for the most part due to a group of bad solder connections, temporarily counteracted the improved memory performance.

The greatest difficulty experienced with the new tubes was due to their centering characteristics. Upon installation a large number of tubes had to be immediately removed, because corners of the 1024 spot raster projected beyond the useful surface of the screen. Ten tubes were removed for this reason. The ORDVAC, like most machines using electrostatic memories, operates its cr tubes in a parallel manner; thus a positioning of the raster which is beneficial to one tube might render the other tubes completely useless.

A test was made on the centering of 40 type C73376B cr tubes and, for comparison purposes, on 40 type 3KP1 tubes chosen at random. The results indicated that the experimental tubes were somewhat inferior in this respect.

In the design of an electrostatic memory, careful consideration should be given to the problem of centering the raster in the cr tubes if the maximum benefit is to be obtained from the unit. It is believed that in a three-inch tube the undeflected beam should be positioned by some method internal or external to the tube to within 2.5 mm of the center of the tube face.

Three tubes have since been removed from the ORDVAC because of the deterioration of the quality of the signals presented to the regeneration amplifier. The signals degenerated to the point that adjustments of the tube voltages would no longer make the signals reliable.

Preventative Maintenance

In September, 1952, a tube-removal program was initiated as a part of a preventative maintenance program. Since almost all of the troubles encountered were the result of shorted or low emission tubes, it was believed that such a program would be helpful in eliminating a potential source of trouble. Blocks of tubes were removed at two-week intervals, and were replaced by new ones. The location and the number of tubes involved were dependent upon the area in which troubles had been occuring most frequently. The number usually varied from 50 to 100. At present the ORDVAC has operated about 18,000 hours, and virtually all of its 3,000 tubes have been replaced at least once. The effect of the

[2] J. M. Wier, "Recent Improvement of ILLIAC Memory," Univ. of Illinois Graduate College, Digital Computer Laboratory, Internal Report no. 45; March 25, 1953.

block-tube change was felt almost immediately, and the procedure is believed to be an important factor in the gradual increase of available time noted in Fig. 1.

Within a period of eight months, 850 tubes were replaced in blocks; out of these tubes, 394 would not pass the inspection given all tubes that are used in the ORDVAC. The tubes used in the computer are, for the most part, types 6J6, 2C51, and 5687. Upon inspection it was found that about 50 per cent of the 6J6 type were bad. The most frequent cause of failure in this type tube was "shorted" elements. "Shorted" is meant to include those tubes whose elements were joined by a high-resistance path so that a leakage current was detected upon inspection. About 30 per cent of the 2C51 type tubes were found to be bad. Again, the most frequent cause of failure was "shorted" elements. The greatest percentage of bad tubes was found among the 5687 type. About 75 per cent of these would not pass inspection. The majority of the rejections in this case was due to low emission and cathode-to-heater leakage. It is interesting to note that the ORDVAC had been operating rather satisfactorily with such poor tubes in use.

The major cause of trouble in the ORDVAC has been tube failures. The conservative design, in which a safety factor of two was used in the rating of components, has made other failures practically nonexistent. Occasionally, however, bad solder connections do appear, as noted before, and they usually appear in groups. Recently, the cause for a great many adder failures was eliminated by the discovery of a wiring error which caused 200 volts to be applied between the heater and cathode of 50 tubes. Such wiring errors are difficult to eliminate except through continued use of the machine.

Perhaps one of the greatest potential sources of trouble in a machine is dust or dirt. In almost any installation some dirt is certain to enter the cooling system regardless of the elaborateness of the filtering system. Simply exposing the components for necessary maintenance will allow a great deal of dust to enter the system. A computer using an electrostatic memory is especially susceptible to dust, unless it is elaborately protected, because the high-voltage wiring forms a fine precipitator. This has been one of the great sources of trouble in the ORDVAC—not great in the sense of occurrence, but great in the sense of damage that may be done.

During the summer of 1952, sufficient dust had collected on the wiring so that the 2,000 volts arced to ground in a number of places. The damage included several clamping tubes that were exploded, at least six-memory chassis that had to be replaced, and five cr tubes that were burned out. The arcing was eliminated by floating the high-voltage system and supplying it with a high impedance variable high voltage. In a darkened room, the arcs were both audible and visible, and they were eliminated by cleaning and separation of the wiring where possible. The manner in which the ORDVAC is constructed makes proper cleaning almost impossible, but a little care has thus far prevented a reoccurrence of the trouble.

Conclusion

Recent improvements to the electrostatic memory have removed certain restrictions on coders and have led to shorter and faster programs.

A preventative maintenance program consisting chiefly of periodic tests and a systematic exchange of tubes, along with the improvements to the ORDVAC, have resulted in an average of 103.10 hours out of a possible 154.46 hours each week over a period of 19 months being made available for computation. It is expected that even a higher number of available hours will be obtained in the future when procedures are perfected and operators gain more experience.

Discussion

Harvey Rosenberg (Burroughs Adding Maching Corporation): How many control relays are used? What was their reliability?

Mr. Williams: In addition to the few relays used to control the power to the ORDVAC there are six telephone-type relays and one stepping relay which are used in conjunction with the teletype output system. No trouble has been experienced with the power control relays; the stepping relay has required minor adjustment only three times in two years; and it has been necessary to replace two of the telephone-type relays because of broken contact arms.

B. B. Paine (Massachusetts Institute of Technology): Could you elaborate on the tube acceptance test program? Is preburning used?

Mr. Williams: Before all tubes are placed in the ORDVAC they are given four tests. The first is a very simple short test in which the tube is tapped manually while a neon is watched. No attempt is made to actually measure the resistance between elements. The second test is an emission test. The third test is a cut-off test. If I remember correctly, in this test about 20 volts bias is placed on the tubes and no more than about 50 microamperes plate current is allowed to flow. The fourth test is a filament-to-cathode leakage test in which no more than 50 ma current at 150 volts is allowed. Tubes are not pre-burned at present, and few tube failures can be credited to this fact.

E. L. Harder (Westinghouse Electric Corp.): Were errors detected in the teletype transmission from the University of Illinois? How many? How detected? How corrected? How was transmission of errors verified?

Mr. Williams: In answer to these questions I will go into more detail on the manner in which the tapes were handled. The University of Illinois initiated contact by transmitting their program directly to the Signal Corps office at Aberdeen Proving Ground, preceeding the program with instructions to the ORDVAC operator. Immediately upon receipt of the tape, it was transmitted back to the University of Illinois where it was compared with the original copy. If an error was detected upon comparison, the procedure was repeated until a correct transmission was verified. In a one month period during the summer of 1952 transmitting time averaged about forty-five minutes a day. There were several human errors made in this period, but there were only two mechanical errors.

Some Remarks on Logical Design and Programming Checks

HERMAN H. GOLDSTINE†

I SHOULD LIKE to discuss two topics which are of some interest. One, is it possible to construct out of elements which are known to be imperfect a logical automaton which will have a given degree of reliability? Two, given a machine, what can one do by way of programming checks to insure accuracy of the final results?

Little has been done on the first problem beyond the primary and extremely important work of von Neumann in this field. This work was presented by him in a series of lectures at the California Institute of Technology in January, 1952. I wish mainly to tell you about this work. Imagine that you are going to construct an instrument out of certain basic organs, such as vacuum tubes, relays, etc. each one of which has a certain probability, ϵ, of failing to perform a function correctly. Can this instrument, or rather logical automaton, be so designed that it will function with a given pre-assigned degree of reliability, i.e. with a probability of failure less than some preassigned value? That is to say, can the mean free path between failures be made as large as you please? Now let me describe briefly the elementary logical building blocks out of which an automaton can be built. These blocks are the tools on which the rest of the discussion is based. Such a construction was shown on paper by the English mathematician Turing.[1]

There are three elements which will be of interest to us in this connection. All these are devices with two inputs and one output. The first is sometimes called a logical "and" element. It has the property that it will give out at a certain time a signal on its output, if and only if signals are present on each input. The second, a logical "or" element, is a device such that if a signal is present on either one or both of these inputs, there will be a response at the output. The third type which is of interest is the logical "and not" element. The output has a response if and only if there is a signal on the first input and not on the second. It can be shown that by means of these organs a logical automaton can be constructed. The remarkable work of Turing was followed with papers by Pitts and McCulloch, in which essentially this type of notational technique was used to study neurological problems. All these discussions, however, assumed that the basic elements had exactly the properties just mentioned. The question that von Neumann raised was the following: Suppose that these elements do not act exactly in the fashion stated, but rather that they have a finite probability ϵ of failing to perform properly. Then, can they still be put together to form, say, a computing machine, which will have a pre-assigned degree of reliability?

To indicate the mode of analysis met with in von Neumann's theory, let us consider an extremely simple situation, a primitive memory cell. We choose a perfect logical "or" element and "feed" the output back in as an input. If this unit is ever excited, it will remain so—we assume the response to come one "time unit" later than the stimulus. Now suppose, instead, that this element has a probability ϵ of failing. Then if p_s is the probability that it will be excited at time unit s, we have

$$p_{s+1} = \epsilon(1 - p_s) + (1 - \epsilon)p_s,$$

or equivalently for small ϵ

$$p_{s+1} - 1/2 = (1 - 2\epsilon)(p_s - 1/2) = (1 - 2\epsilon)^s(p_0 - 1/2) \leqq e^{-2\epsilon s}(p_0 - 1/2).$$

Thus as s increases, p_{s+1} approaches 1/2, i.e. it becomes equally likely that the unit will or will not be excited, independently of the original state. In spite of this pessimistic picture it is possible to arrange elements to prevent such degradation of information or performance, as we shall see.

Von Neumann shows that the basic blocks described above may be replaced with a single one. Consider his "majority" organ. It is a three-input unit whose output responds if and only if at least two inputs are stimulated. Now suppose that the probability of failure of this organ is ϵ, and the probabilities of wrong signals on the inputs are η_1, η_2, η_3, respectively. The question now is, "What is the probability that an error signal will appear on the output?" A gross over-estimate for this probability is $\eta_1+\eta_2+\eta_3+\epsilon$. Since many of these organs are to be used in cascade, it might seem that the probabilities of error would cumulate enough to prevent the desired construction. Fortunately one can make certain assumptions about the way the organ is to be used. It can be shown, using sharper estimates, that under these assumptions the error does not grow beyond bounds. Suppose that the probabilities of errors on the input lines are indpendent, and that under the proper functioning of the network these lines are always in the same state of excitation: the lines are either all stimulated or all not stimulated. This latter assumption seems quite restrictive; however, it can be shown that in building an automaton almost all the organs used can have this property. So, under these assumptions, it may be shown that the probability θ of failure of a perfect organ, due to input errors, is

$$\theta = \eta_1\eta_2 + \eta_2\eta_3 + \eta_1\eta_3 - 2\eta_1\eta_2\eta_3.$$

† The Institute for Advanced Study, Princeton, N. J.

[1] *Proc. London Math. Soc.*, vol. 42, pp. 230–265; 1937.

This is the probability that at least two of the wires will be wrongly stimulated. The probability of failure of an imperfect organ is

$$\bar{\eta} = (1 - \epsilon)\theta + \epsilon(1 - \theta).$$

If we replace the numbers η_1, η_2, η_3 by an overbound η, we have

$$\bar{\eta} = \epsilon + (1 - 2\epsilon)(3\eta^2 - 2\eta^3).$$

Can this number be itself smaller than η? It will be shown below that the answer is, under certain circumstances, yes. The fact that we can make $\bar{\eta}$ not greater than η means that the output of this organ will have a probability of failure which is not worse than the probability of failure of any of the input lines; it suggests the possibility that one could build an organ which would not be progressively degenerative in character. This, among other things, is what von Neumann has shown. Oneway to do it is the following: Suppose that the automaton to be constructed is a "black box" with many inputs. Let us make this box in triplicate and parallel the inputs. We have triplicated the size of the system, and we are now going to ask for a majority vote. That is to say, we will examine the outputs of these three units and take as correct the output of two out of the three. If η is an upper bound for the probability of error at any input of any one of these three organs, then the probability of error at the output is the $\bar{\eta}$ derived above. How does $\bar{\eta}$ depend on η? We consider the function $\bar{\eta}$ of η on the interval 0 to 1. For $\epsilon \geqq 1/6$, we have $\bar{\eta}$ nearer to 1/2 than η is. Thus in a cascade of such elements with $\epsilon \geqq 1/6$, $\bar{\eta}$ will tend to 1/2, i.e. the system will degrade to a state of indifference. We must therefore suppose $\epsilon < 1/6$. In this case we shall see that $\bar{\eta}$ will tend to $\eta_0 = \epsilon + 3\epsilon^2 + \cdots$. For $\epsilon < 1/6$, the curve $\eta^* = \bar{\eta}\ (\eta)$ intersects the line $\eta^* = \eta$ in three points:

$$\eta^* = 1/2, \qquad \eta^* = \eta_0 = (1/2)\left(1 - \sqrt{\frac{1 - 6\epsilon}{1 - 2\epsilon}}\right),$$

$$\eta^* = (1/2)\left(1 + \sqrt{\frac{1 - 6\epsilon}{1 - 2\epsilon}}\right).$$

In the interval $(0, \eta_0)$ $\eta \leqq \bar{\eta} \leqq \eta_0$; in $(\eta_0, 1/2)$, $\eta_0 \leqq \bar{\eta} \leqq \eta$; thus in cascade $\bar{\eta}$ will tend to η_0.

This heuristic argument suggests the possibility of constructing an automaton with probability $\cong \epsilon + 3\epsilon^2$ of failure, out of basic elements with individual probabilities of failure ϵ, provided that ϵ is not more than 16 per cent. A rigorous argument has been made by von Neumann. It shows that the total number of organs needed is of the order 3^μ, where μ is a measure of the number of organs to be cascaded. A value of μ of about 160, which is not unreasonable, gives $3^\mu \sim 2 \times 10^{76}$; this is more than the total number of electrons. This value results from triplicating rigorously, as required by von Neumann's construction. It will be of great interest to see whether substantially more favorable estimates can be obtained by modifications of his construction, and whether significant improvements would obtain by alterations in the performance requirements.

He also has attacked the problem along quite different lines. This involves multiplexing of the lines carrying information. Each majority organ of the original automaton is to be replaced by a new organ having bundles of N wires instead of single ones for inputs and outputs An input is regarded as stimulated if at least a certain number, say $(1-\delta)N$, of the wires of a bundle are stimulated, and as not stimulated if at most $\delta\ N$ are. If a number between these is stimulated, an error is regarded as having occurred.

To construct an automaton using bundles instead of single wires it is desirable to use two types of organs, restoring and majority. The former, is as the name suggests, used to overcome degradation of information content of a bundle by causing all output wires to be stimulated or not, depending on the number of stimulated input wires, according to the definition above. (A trivial example of a restoring organ is a majority organ whose output is triplicated, in this sense: if we think of the three inputs as being a bundle, then none or one of the wires of the bundle stimulated produces zero stimulated at the output, whereas two or three stimulated produces three.) It is not difficult to show how a majority organ in this system can be built out of those we have discussed previously.

Von Neumann shows that $\delta = 0.07$ is in some sense most favorable. For this value, and for ϵ of not less than 0.0107, it is not possible to improve the reliability of the system beyond a certain point no matter how large N may be; he accordingly chooses $\epsilon = 0.005$. Then for 1,000 lines in a bundle the probability of malfunction is 2.7×10^{-2}; for 10,000, 1.6×10^{-10}; and for 20,000, 2.8×10^{-19}.

Let me now say a few words on programming checks into computations in order to monitor the accuracy of calculated results. (The succeeding speaker will, I understand, discuss checking by circuit techniques.) The mathematical operator has two tasks. He has to insure first that the initial code is correct, and second, that the machine does not malfunction during the operation of the problem. The debugging phase of the problem and detection of errors in the machine are interrelated and may be treated similarly. During the debugging phase the operator is interested in finding out whether the code reflects precisely the mathematical algorithm. This task appears different from watching the machine for errors, but in reality it is not. An error either of the code or of the machine itself is to the operator a failure to produce a correct result. So he can act as if both situations were the same, and prepare his techniques accordingly. These center principally around redundancy techniques. By that I mean the following: in order to find out whether the machine is proceeding correctly, whether this be a trouble in the code or in the machine, the mathematician can introduce into the problem certain conditions which must be satisfied if the calculation is proceeding correctly. Let me give a simple example. Suppose that the calculation consists of the numerical integration of the differential equations describing a conservative system, i.e. one in which the total energy must be conserved. If a periodic calculation of the total energy gives agreement with the initial energy, the operator has a fair assurance both that his code is correct and that the machine is not malfunctioning. One can, in a great many problems find some rather deep redundancy which will serve as such a check.

Discussion

M. Stateman (*Sylvania Electric Co.*): Would you give the specific references and availability of pertinent papers by von Neumann and Pitts?

H. H. Goldstine: I can give references to von Neumann's paper, Pitts, McCulloch, Turing, and one other paper which is also relevant. The fundamental paper is by von Neumann. It consisted of a set of lectures delivered by him. The title of the paper was "Lectures on Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components." These were lectures delivered by von Neumann at the California Institute of Technology, January 4–15, 1952, and they were prepared from notes by R. S. Pierce. The W. S. McCulloch and W. Pitts papers are in the *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115–133; 1943. The paper of Turing's, which is the first paper I know of on the subject, is in the *Proceedings of the London Mathematical Society*, vol. 42, pp. 230–265; 1937. There is also a paper by Kleene which is a Rand memorandum.

Kenneth Rose (*Burroughs Adding Machine Corporation*): In connection with the figures given on the effectiveness of multiplexing, it should be pointed out that von Neumann assumed a probability of error of 0.005 in each component. Such unreliability of course affects the results.

Mr. Goldstine: This is perfectly true. In discussing the results on multiplexing, I forgot to make mention of the following facts. In getting the final estimates, von Neumann made the assumption that the number delta was 0.07. That is to say that a signal of one sort ("Yes" or "No") would mean that at least 93 per cent of the wires were stimulated and the other sort not more than 7 per cent, the error band lying in between. Having made the assumption of delta of this size, the epsilon then comes out to 0.005, or a half of 1 per cent reliability.

H. Campaigne (*National Security Agency*): You said that with 1,000 wires reliability of 2.7×10^{-2} could be achieved. Would you repeat that statement emphasizing the conditions under which it holds?

Mr. Goldstine: The statement is essentially of this sort. In making final error estimates, one has a probability integral, the size of which he estimates. If you tabulate the number of lines in a bundle against the probability of malfunction, then at the low end, say for 1,000, the probability is 2.7×10^{-2}, and at the high end, say 25,000, this is 1.2×10^{-23}. Now to indicate what this means, let me consider an example. Suppose that you want to operate a machine, a computing machine, with 2,500 vacuum tubes. Suppose furthermore that each vacuum tube is actuated once in every five microseconds. Suppose you want to run with a mean free path of eight hours between errors. This means that you will perform about 1.4×10^{13} actuations, which means then that if you take the tabular data you find that you want a multiplexing of about 14,000 wires. Again comparable statements could be made about the number 2.7×10^{-2}. You see that in order to evaluate the probability of malfunction one has also to measure the speed of operation of the machine and the mean free path that one desires between errors.

C. Huang (*Sylvania Electric Products*): Would you please give a general comment on the relationship between η^*, η and ϵ, and how to minimize η^*?

Mr. Goldstine: That formula, as you recall, was not from the multiplex case. It was in the other case of attempting to achieve a reliable arrangement by taking a majority of three. The formula there was of this sort:

$$\eta^* = \epsilon + (1 - 2\epsilon)(3\eta^2 - 2\eta^3).$$

In the first place we indicated that ϵ wanted to be less than 1/6 in order to get any reliability at all. The number ϵ was the probability that a basic organ would malfunction. η was an overbound for the probability that any one of the input leads to any one of the basic organs would carry the wrong information. η^* was the probability that the output of a given organ would malfunction, that is, the output would be erroneous. You cannot exactly minimize η^* because what you are interested in is a sequence of these organs, one feeding to the next. What you have to do is this: You start with a given organ with a given probability η coming in. The output has a probability η^* that is fed into a second organ. You then get another η^* out of the second organ. The problem that is relevant here is, "What in effect is the limiting or the asymptotic value of this?" In other words, "If you operate a very long sequence of organs, entering with errors bounded by η, what is the probability of error at the final output?" I hope this answers this question.

Ralph B. Conn (*California Technical Jet Propulsion Laboratory*): I would like to make a comment concerning the availability of copies of notes from von Neumann's lecture: There are none. There were about 200 to 300 people that attended the lecture, and practically everyone who attended the lecture got a copy of the notes. If you know anyone out there, I suggest you write them and try to borrow a copy. That is probably the best way to get a set.

The Advantages of Built-in Checking

JOHN W. MAUCHLY†

ELECTRONIC COMPUTERS save money. That is the basic economic fact behind all of the current activity in the designing and building of information-processing systems. The job of the engineering designer is economic as well as technical. He must balance many factors as well as he knows how in order to achieve maximum utility at minimum cost.

From this simplified point of view, it might seem that all the potential user needs to do is to state his problem and let the engineer design the appropriate equipment. This concept of engineering design leads us to the term "black box." Every engineer is familiar with the black box. It transforms any input into any output. Mr. X, the user, specifies his input information and the output which he desires. The job of the engineer is to produce the black box which will do this. No one but the designer need concern himself as to what is in the box.

However, it turns out that Mr. X wants a black box which the engineer cannot make. Mr. X would like perfection. He wants a system which never breaks down, and never makes errors. Although progress is being made in the improvement of computer reliability, no one is yet ready to guarantee his computer perpetually perfect.

For this dilemma, the engineer has a handy solution. It sometimes goes under the name of cybernetics. The engineer calls it "feedback." The accountant calls it control. The idea is to approximate the desired perfection by measuring, so far as we can, the deviations from perfection and using such information to get a better approximation. Of course, it must be understood that we shall never attain the perfection which is our goal, but we may approximate it as closely as we please. It is again an economic problem as to how much error we are willing to tolerate, and how much it will cost to reduce the errors so that they are within this tolerance.

For any feedback method to work, there must first be a method of detecting the errors which are made, and there must then be a method of correcting these errors. There are numerous schemes for doing these things. I shall pick a few representative methods for enumeration. (1) The computer checks each individual arithmetic operation or transfer in a way which is likely to catch almost all errors and, when an error is detected, the computer repairs itself and continues it work. (2) The computer automatically detects almost all errors and, when an error is detected, it tries the same computation again. If no error is made on the new trial, it continues its work. If the error persists, the computer continues its attempts. (An automatic circuit might indicate that repair is needed.) (3) The computer has automatic error-detection circuits which cause the computer to stop, at which time an operator takes appropriate action. (4) Every problem is run at least twice, and diagnostic tests are introduced at intervals. The diagnostic tests verify that the computer is operating correctly at the time the test is run. Results from two different runs of the same problem are compared, and agreement is taken to indicate no error. When, by diagnostic tests, or disagreement of problem runs, a malfunction is indicated, appropriate maintenance action and problem-result correction is taken. (5) The computer has no automatic checking, and complete reliance is placed upon programmed checking to insure correctness of problem. Diagnostic tests or other means are used to indicate correct operation of computer.

It should be noted that marginal checking is included in the term "diagnostic tests." Marginal checking is a method for increasing the reliability of operation, not a method of verifying or checking problem results.

There are numerous variations and combinations of checking techniques which might be enlarged upon, but the above list will suffice for this discussion. We will leave the discussion of self-repairing computers to the future and, particularly, to the meetings in New York this March. Equipment of this sort has been built, but it is not commercially available. The Remington Rand electronic card calculator operates according to the *second* scheme. The UNIVAC is more like the *third*, since it contains error-detecting circuits which stop the computer, and the operator must then take appropriate action. The *fourth* scheme, that of running the problem twice, has been a favorite method for use with all sorts of computing devices, including punched-card machines. It is sometimes advocated as a reasonable procedure for large automatic electronic digital equipment, although it is clear that this scheme cuts machine efficiency in half, at the very least. In other words, the operating cost is more than doubled when every problem is run twice. The *fifth* method, that of using programmed checks, is a much more difficult one to evaluate.

In a paper given last March entitled "Checking Circuits and Diagnostic Routines," J. P. Eckert discussed the features of the ENIAC, the BINAC, and the UNIVAC with regard to checking methods and error-detection facilities. He estimated that the error-detection circuits and the duplication necessary to provide an adequate check on UNIVAC operations amounted to less than 30 per cent of the cost of the central computer and, therefore, an even smaller percentage of the

† Remington Rand Inc., Eckert-Mauchly Div., Philadelphia, Pa.

cost of the entire computing system. It should therefore be obvious that the duplicate-run procedure for screening out errors is certainly at an economic disadvantage. Such a procedure has the effect of doubling the cost of the computer, rather than adding only 25 per cent or 30 per cent to its cost.

What can be said about the effect of programmed checking on the cost of computer operation? Is programmed checking an economic substitute for built-in checking? It should be clear that there is no universal answer to these questions. The cost of programmed checking will vary with the application. The efficacy of such methods must also vary with the application. It may well be that when a computer is to be used day-in and day-out on problems which are, in the main, of a mathematical nature, the cost of programmed checking can be much lower than the 25 per cent or 30 per cent which we have charged to built-in checking. So much depends upon the field of application of the computer that no general statement regarding the cost of programmed checking can be made.

Probably the most important aspect of programmed checking is that the feedback loop discussed above occurs entirely *outside* of the black box which the engineer has designed. Hence, the engineer is tempted to disregard this problem as one which is, by definition, outside of his province. The user has no interest in the cost of results which are of unknown accuracy, but is vitally interested in the cost of the final results *after* the feedback process has screened out the incorrect runs and reduced the chance of error to an acceptable minimum. The engineer would like, in all honesty, to design a computing system which gives the user just what he wants at the lowest cost. The engineer has, however, no control over the external feedback loop which is dependent upon programmed checking. This also means that he usually has very little information about the properties of the over-all system which includes such a feedback loop. Without a great deal of information on this topic, it is obviously rather hopeless for the engineer to judge the merits of a programmed checking system versus an automatic built-in error-detection system.

We see then that some part of the difficulty of evaluating the merits of built-in checking versus the merits of programmed checking lies in the very fact that the automatic checking features of any computer can be well defined and their costs estimated as closely as is needed for any comparison, while programmed checking is, by its nature, variable from problem to problem, and from application to application, so that no simple statement can be made with regard to its cost.

These different methods of controlling errors not only affect the cost of the equipment and the running time during proper operation; they also interact with the maintenance and repair of the computing system. This makes the evaluation of their relative merit even more difficult. Built-in error-detection circuits prevent errors from propagating further by stopping the computer at once. They also help to locate the fault without resorting to a separate diagnostic test. The halting of computer operations at the exact instant that a fault occurs is often extremely helpful to maintenance personnel in quickly discovering the exact nature of the malfunction.

A further difficulty in the evaluation of the cost of programmed checking is that there are a multitude of ways in which programmed checking may be employed on some problems, and almost no way by which others can be program checked. If the programmed checking is extremely detailed, and each arithmetic operation or transfer is separately checked, then programmed checking doubles the problem running time and also increases the labor of programming. A programmed check of this nature is obviously not an economic competitor of built-in checking.

However, there are certain types of problems, particularly those of a highly mathematical nature, in which a rather lengthy computation may be checked by a much smaller amount of calculation. When the running time added to the problem in this way is less than 30 per cent, and the programming effort required to incorporate the checks is trivial, then programmed checking may be the economic solution. It is necessary, however, for this criterion to be satisfied for almost every problem (or at least for a good majority of the problems) to be run on a given computer before one can feel free to omit built-in checking.

The point just made leads directly to the justification for including built-in checking as a part of the UNIVAC system. The UNIVAC system is intended to be a general-purpose computer and, in its design, no assumptions could be made regarding the type of problem which would be run on this equipment. There was no way of knowing in advance what difficulties might arise in trying to work out programmed checks for a diversity of commercial and statistical problems. In this connection, it should be pointed out that almost every computer is ultimately called upon to handle problems which are far different from those which were originally considered at the time that the computer was designed.

In conclusion, I would like to point out that there may be sound reasons for adopting programmed checking for some phases of computer operation and automatic error detection for other phases of operation of the same computer. The inherent reliabilities of the various components and subassemblies of the computer are not necessarily all identical.

It is worth noting that, in spite of the rather complete built-in automatic-checking features of the UNIVAC system, programmed checking is used to guard against human errors. One example will suffice. In problems which run for a considerable length of time and use many tapes in the course of the problem, much computing time might be lost through interchanging two tapes, or placing an entirely wrong tape on a particular servo. It is therefore good practice to incorporate in the programming a system of tape-labeling. The UNIVAC then checks, at every step of the problem where a new tape is used, whether the correct tape has been mounted. It is difficult to see how this problem could be met with sufficient flexibility by a built-in automatic checking device.

Built-in checking represents a fixed additional cost which can be estimated rather closely. Programmed checking is not easy to define and, in some cases, is difficult to secure. For a general-purpose computer, the cost of programmed checking is almost an unknown quantity. Reliance on programmed checking results in an unknown and variable operating cost rather than a fixed known initial cost.

Discussion

George J. Christy (State of California, Department of Education): Has there been any major problem in maintaining check circuits in operation?

Dr. Mauchly: The answer is no. I think that is a fair answer even though it is brief.

Dorothy T. Blum (Department of Defense): What provision, if any, is made in self-checking computers to distinguish between errors in the central computer and errors in the checking system? What is the experience on UNIVAC, for instance, on the proportion of error in the central computer to the errors in the checking circuits?

Paul Ellinger (Department of Defense): What possibilities are there for checking the built-in checking circuits themselves?

Dr. Mauchly: I thought these two questions could be handled together with a little further reference to the equipment which the UNIVAC has with regard to checking. You can find more information, of course, in the Eckert paper published last March, and still more by requesting such information from Remington Rand. The 30 per cent increase in the cost of the computer attributed to built-in checking should not be taken to imply that you can segregate 30 per cent of the computer from the rest of it and say that this 30 per cent consists of checking circuits. Different parts of the computer are checked in different ways and in particular, outside of the arithmetic operations, the watch dogs against error are almost all based upon what we call a parity check, that is, using a redundant code and checking that an odd number of pulses are always to be found in the binary digit code which represents *each computer digit.* Parity checks are used throughout the system, not just on the central computer. This includes all tapes and auxiliary apparatus. Also, in the central computer, the acoustic storage tanks carry computer digit information in this redundant code. Hence, part of this 30 per cent we are talking about results from the fact that the memory tanks have to be slightly larger, because the redundant code takes up about 15 per cent more memory.

The checking circuits in the arithmetic part, however, do depend on duplication. There are duplicate arithmetic units, duplicate adders, duplicate registers for temporarily holding the information used in the computation, and so forth. Even some of the checking circuits are duplicated.

There is also some redundancy of checking. There are cases where the results are checked in more than one way. Furthermore, if one parity check circuit fails to catch an error, some other parity check will catch this error an instant later. Then, in examining what went wrong, the finger would be put on the offending elements.

Even back in 1945 and 46 when we were first talking about acoustic computers—Dr. von Neumann was with us in some of those discussions—we were entering into an elementary discussion of some of the problems that Dr. Goldstine just talked about. The question was, "What checks the checker?", just as was asked here, and, "Do you have to have an infinite number of checkers in order to check the checkers, or is this a finite series which you can terminate and still get something reasonably practical out of it? The answer is that you can, without going into an infinite sequence of checking operations, have your output as free from error as you choose, depending on the economic considerations governing the design.

Relative to the question, "What is the experience on the proportion of error in the central computer and in the check circuits?", we see that there is no simple way to allocate the errors to these two seemingly different parts of the computer because they aren't so easily divorced from each other in the actual apparatus.

Joseph Weinstein (Signal Corps Laboratories): Would real-time data processing problems (for control of operations) fit into the category of problems for which program checking is desirable, or would you consider program checking incompatible with real-time operations?

Dr. Mauchly: Here again there isn't any one answer to this. I would hesitate to say dogmatically that it is incompatible. When you're faced with a real-time problem, it is sometimes not solely an economic problem. You may get problems in which human life, or some safety considerations, are involved. You can't measure these in dollars and cents, so you may have to multiplex exceedingly in order to prevent errors.

Edward H. Friend (U. S. Navy): Can you indicate the degree of progress attained in automatic coding research for UNIVAC and other computers?

Dr. Mauchly: There are various things called automatic coding, ranging from translations and pseudo-codes through compilers, generators, and very fancy kinds of automatic coding. In our own work very good progress has been made on problems of a mathematical nature, but we're not quite so far along on problems of a commercial nature, but they're all receiving attention. If you want more information on this, there are UNIVAC-using agencies right here in the area where such methods are used, and some of these centers have been the scene of full-day work-shops on automatic coding. There is quite a lot of information available, and I am sure that Dr. Grace Hopper will be glad to discuss her work in this field with anyone who is interested.

Recent Progress in the Production of Error-Free Magnetic Computer Tape

J. C. CHAPMAN† AND W. W. WETZEL†

Introduction

SIGNAL DROPOUTS arising from magnetic tape are one cause of error in modern digital computers designed to use such tape as a long period storage medium. Noise pulses, which are of sufficient amplitude to act as spurious signals, form a second error source. Both dropouts and noise pulses are traceable to discontinuities in the magnetic coating.

During the past year extensive studies have been made of coating defects with the intention of minimizing their occurrence. We initially believed that if the discontinuities could be examined and classified there was hope for improving tapes by eliminating the defects at their source. This has been done and the practical results are gratifying.

This paper will explain briefly: (a) our findings concerning the physical causes of errors; (b) the reasons why errors arise from such physical defects; (c) steps taken to eliminate errors; and (d) a summary of our progress during 1953.

Detection of Errors

The equipment used for dropout detection records square-wave pulses at a rate of 100 to the inch on seven tracks of a half-inch wide tape. The recorded tape is read back and if the signal for any pulse falls below 55 per cent of the normal maximum value, an error is recorded. At each error the tape is stopped automatically so its physical cause can be examined under a microscope.

Noise pulses which are counted as errors are located by saturating the tape continuously in one direction and stopping the tape during read-back when a noise peak exceeds 8 per cent of the normal maximum signal. Again, microscopic examination is used to find the physical cause of the noise pulse. Generally speaking, dropouts and noise pulses arise from the same physical causes and they have been grouped together under the general name of errors.

We have further classified errors as removable and nonremovable. Removable errors arise from loose particle contamination of the coated surface and may usually be cleaned off the tape with a soft brush. Nonremovable errors may be caused by oxide clumps or foreign particles which are embedded in the coating. Since removable errors may be eliminated through inspection they are at present considered to be unimportant. The data which follow will be confined to nonremovable errors.

† Minnesota Mining and Manufacturing Co., St. Paul, Minn.

Causes of Errors

Nonremovable signal dropouts may be caused by a lack of magnetic coating at the point where a pulse is supposed to be recorded, but the early experience of observers led to an explanation based on the more frequent occurrence of small inclusions (called "nodules") in the coating. Upon close inspection, these nodules could be classified as oxide clumps, acetate particles, embedded filter fibers, etc. Initially oxide clumps were the most frequent offenders.

Oxide clumps which protrude from the otherwise flat surface of the tape force the main body of the tape away from the recording and playback head gaps. During recording the effect of the presence of a nodule is to reduce the sharpness and the intensity of the recording field at the tape surface. On playback, where the rate of change of recorded flux is observed, the already reduced steepness of the flux front is observed from a distance which further reduces the rate of change of flux in the reproducing head. This combination results in a decrease in output which is called a dropout. If the dropout is sufficiently large it constitutes an error.

Noise errors similarly arise from discontinuities in the magnetic coating. The tape is magnetized to saturation longitudinally prior to read out in the noise test. The flux seen by the playback head under this condition of magnetization would be essentially zero for a perfect tape except for small variations in leakage flux from particle to particle of the oxide which causes the normal noise background. However, if a discontinuity in the coating occurs, which might result from a pin hole or the inclusion of a particle of nonmagnetic contaminate, magnetic poles will form on either edge of the gross discontinuity and the leakage flux will increase well above that due to the normal physical separation between oxide particles. A noise pulse also arises from nodules of oxide where excess magnetic material is present.

Elimination of Errors

As a preliminary step to elimination of errors from computing tapes, a study was made of the error types and their frequency of occurrence. It seemed obvious that if the causes of errors were known we would have a clue to the step in the tape-making process where they were introduced.

One typical example of oxide clumps is shown in Fig. 1. When such clumps are encountered during playback, they give rise to both dropouts and noise errors providing the clumps are of sufficient size. Fig. 2 shows a coating streak where absence of oxide caused

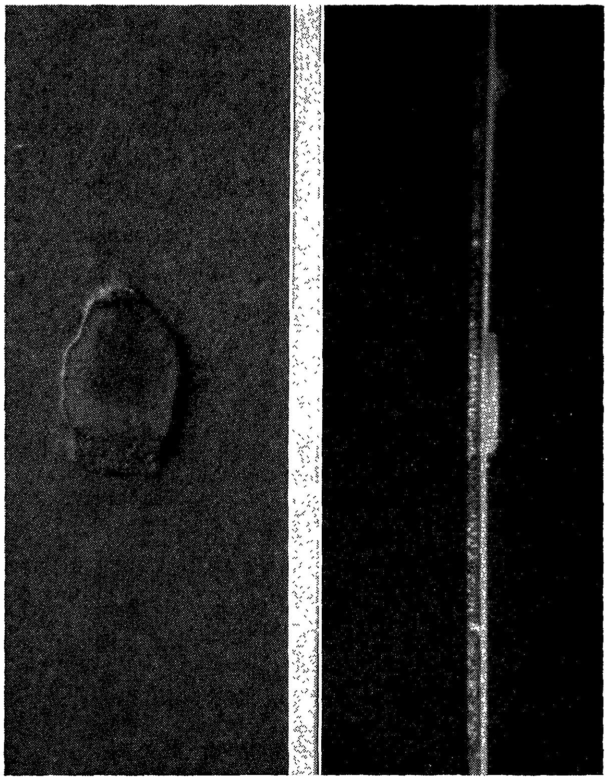


Fig. 1—Showing oxide flake embedded in coating and cross section of same (magnified 50 times).

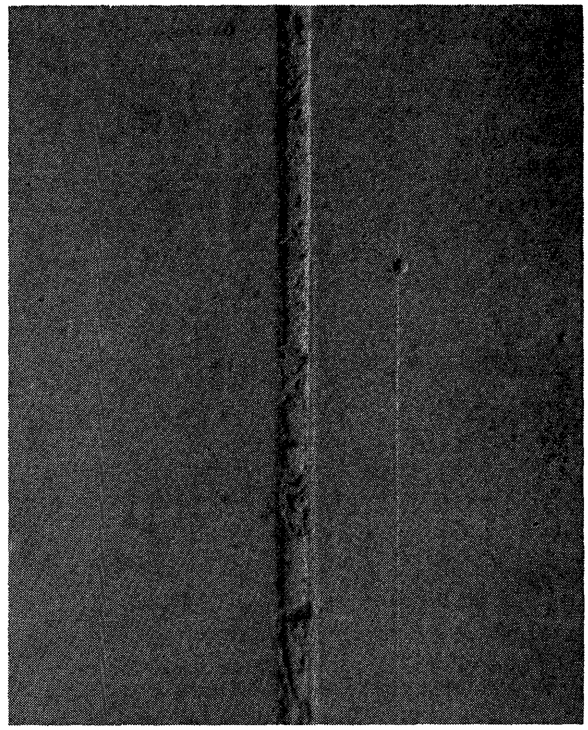


Fig. 2—Showing a streak in the magnetic coating which results in a deficiency of magnetic oxide along the streak (magnified 50 times).

both dropout and noise errors. Fig. 3 illustrates our findings in March, April, and May on the frequency and cause of nonremovable errors. Each sample represents 24 rolls of ½ inch×2,400 foot tape.

It will be seen that oxide flakes, tape distortion, and acetate particles were the predominate sources of error in March. A test run in April, 1953, was designed to eliminate oxide flakes, filter fibers, and tape distortion. Tape distortion arises primarily from creases in the tape due to faulty handling during the manufacturing process. The results of the initial experiments are shown in Fig. 3. We were successful in reducing to zero the number of errors from the three sources being studied.

Two runs were made in May in which special precautions were taken to eliminate acetate particles. We were apparently on the right track since both acetate and miscellaneous embedded particles were reduced in frequency of occurrence. The increase in oxide flakes to 1.5 errors on an average pointed out that our April observation was possibly based on too small a sample.

Since better than 50 per cent of the rolls during the last two trial runs were error-free, we decided to turn the process over to production. Fig. 4 shows the error count on production runs of relatively large samples. In order to compare the results with previous test runs, the bar graphs have been reduced to 24-roll equivalents. The runs illustrated total 664 rolls, ½ inch×2,400 feet.

In changing to production, our classification of errors was refined. Certain defects previously entered under miscellaneous particles were found to be attributable to imperfect backing. The acetate film used as tape back-

CAUSES OF ERRORS	MARCH	APRIL	MAY	MAY
OXIDE FLAKES	20	0	1	2
ACETATE PART.	15	10	2	4
FILTER FIBERS	8	0	0	0
STREAKS	8	9	0	0
MISC. PARTICLES	8	7	4	4
DISTORTION	19	0	0	0

■ NUMBER OF ERRORS

Fig. 3—Chart shows number and causes of errors found in four experimental lots of 24 rolls, each of ½-inch×2,400-foot computer tape. (In May experiments much improvement has been shown over March.)

CAUSES OF ERRORS	JULY 36 ROLLS	AUG 252 ROLLS	AUG 76 ROLLS	AUG 78 ROLLS	AUG 222 ROLLS
BACKING	4.0	4.1	1.6	1.5	1.6
DISTORTION	3.3	5.3	0.3	2.8	2.5
PINHOLES	0	2.6	0	0	0
FILTER FIBER	0	0	0	0	0.1
STREAKS	0	0	0	0	0.1

■ NUMBER OF ERRORS FOUND IN A 24 ROLL SAMPLE

Fig. 4—Chart shows number of errors and rolls in certain production runs of computer tape. For comparison with Fig. 3 number of errors has been reduced to equivalent found in a 24-roll sample.

ing is cast on large diameter wheels having a polished surface. The film began to show tiny defects which were determined to be repetitive and which could therefore be attributed to a small dent in the surface of the wheel. The errors due to this source are listed as backing defects. Fig. 5 shows one such defect which occurred repeatedly in the acetate film.

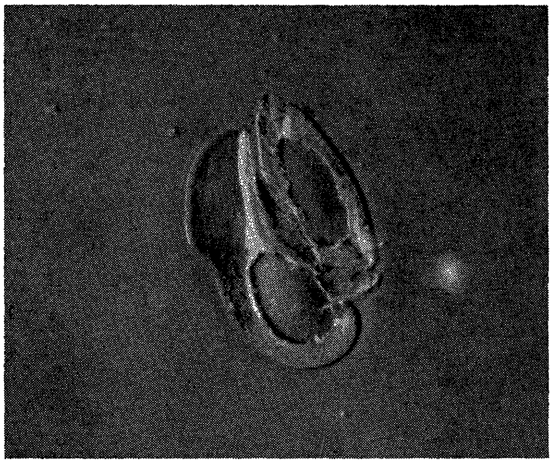


Fig. 5—This photomicrograph (magnified 25 times) illustrates appearance of a defect in acetate backing of magnetic tape. This irregularity is reflected in magnetic coating and gives rise to an error.

Tape distortion which arises from the permanent creases in tape backing occurs most frequently on the edge of the tape. The errors which arise from pin holes, from filter fibers, and from streaks, now appear to be purely random.

The bar graphs show that, through studies of errors and through efforts made to eliminate them, we have managed to reduce errors from 3.25 per roll in March, to 0.18 per roll in August.

Conclusion

To a tape manufacturer our studies show that if special precautions and techniques are used, production waste figures can be held within reasonable limits during runs of computing tape.

To the consumer of tapes it means that error-free tapes are available, providing each tape is individually checked by the manufacturer. Alternatively, if the consumer prefers to check each roll of untested computing tape, he may expect better than 75 per cent of the rolls to be error-free.

Discussion

F. Hawkins (David Taylor Model Basin): Is there any deterioration of the magnetic tape record with time? If so, what are the causes?

Mr. Wetzel: As far as we are able to determine the only deterioration of magnetic tape records comes from mechanical failure. Nicks in the edge of an acetate tape will cause tears. If the tape is run too often—I don't know how much that would be, 25,000 times or something like that–then the oxide will tend to abrade off. It is my opinion that the ultimate failure will be mechanical brought about through nicking of tape edges.

Mr. Hawkins: What is the frequency of errors to be expected on the best magnetic tape presently marketed by your company?

Mr. Wetzel: We are marketing No. 109, a so-called instrumentation tape. The data which I presented today gives the most recent figures we have on the frequency of errors in this tape. The figure which I gave in the summary is 0.18 errors per ½ inch tape, 2,400 feet long.

David Rutman (Rand Corporation): What width of recording track do you use in the tests?

Mr. Wetzel: The track width as I remember is 35-thousandths of an inch, give or take five mils for a poor memory.

Mr. Rutman: What is the size of the imperfections?

Mr. Wetzel: The size of the most frequent imperfections is of the order of 10-thousandths of an inch in diameter. The long streak I showed you on the board is something in the order of 1/10 of an inch long and five or ten mils wide.

Mr. Rutman: When will Mylar backing be available?

Mr. Wetzel: A member of the Dupont Corporation would have to answer that, really. Our best crystal ball, which is somewhat clouded in this prediction, says that Mylar backing for computer tape should be available some time in 1956. The Dupont people are quite certain that their plant will be in operation the latter part of next year, but I am afraid the perfection of film required for computer applications will not be attained for some additional time. If you recall, one of the most frequent imperfections in acetate backed tape which we observed in the last test runs are caused by dents in the casting wheel. These tiny defects will show up as errors. I think Mylar which now contains many inclusions is quite a way from the degree of perfection which acetate film has achieved after 25 to 30 years.

C. P. Bastuschek (Haller Raymond and Brown, Inc.): Has any check been made on the number of errors per reel of the aluminized tape?

Mr. Wetzel: Aluminized tape as it is generally delivered is used for static elimination. The aluminized tape is not the No. 109 computer tape quality, and I would guess—it would have to be a sheer guess—that something in the order of 10 errors per reel could be expected. It is not the computer quality material.

J. A. O'Brien (M.I.T.): Can you give information on Mylar tapes?

Mr. Wetzel: Mylar is a very intriguing material since it has relatively high tensile strength and very excellent tear resistance. It stands both higher and lower temperatures than acetate. The one reason that we do not recommend its use in computers is that the film at present has a good many defects in it. The tensile strength of acetate compared with Mylar is not too terrifically different. One-and-a-half mil acetate has the same strength as one mil Mylar. The superior tear resistance of Mylar is the outstanding quality you may expect under normal operating temperatures and humidities.

Reliability of Electrolytic Capacitors in Computers

MARK VanBUSKIRK†

IF ANY PERSON would ask ten different electronic manufacturers whose equipments cover a wide variety of electronic circuits for their opinions of electrolytic capacitors, it is possible there would be ten different answers. These answers might range from: "We use them all the time, and never have any trouble," to the extreme: "We wouldn't use them under any circumstance. They are not reliable."

Obviously, the first manufacturer has capacitor applications in which electrolytic capacitors operate satisfactorily. Also, the other manufacturer has no capacitor application in which electrolytic capacitors will operate satisfactorily and probably he has tried to use electrolytics with completely unsatisfactory results.

Which of these attitudes applies to electrolytic capacitors in computers?

You want to know whether electrolytic capacitors will operate reliably in computers or not. If they will, you want to know what procedure must be followed to insure reliable operation. Rather than to list specific applications in which electrolytic capacitors will operate reliably, it will be best to present the information by which a designer may determine in what applications and under what conditions electrolytic capacitors will give reliable operation.

Speaking of the reliability of components, there is no question that any component can be damaged by attempting to operate it under too severe conditions. Thus the point of deliberate misapplication need not be considered.

The first point to consider about the reliability of electrolytic capacitors is their theoretical reliability. If this is not good, there is no need to consider the problem beyond this point. Fortunately for the electrolytic capacitor industry, this theoretical reliability is excellent. The necessary elements of any capacitor, two conductors separated by a dielectric, are, in the electrolytic capacitor, materials with no inherent weaknesses. Also, the materials are such that if a flaw develops in the dielectric, the voltage applied to the capacitor will cause the flaw in the dielectric to be eliminated automatically. Another factor which indicates good theoretical reliability is that there are no moving parts in electrolytic capacitors.

Having determined that the theoretical reliability of electrolytic capacitors is excellent, the next step is to look at the practical side of the picture. Laboratory and field experience verify the conclusion of theoretically good reliability. Fig. 1 shows the electrical characteristics of an electrolytic capacitor after more than four and one-half years of accelerated laboratory life test operating at 65 degrees C. and rated voltage. All electrical characteristics are desirable. Capacity is greater than nominal, equivalent series resistance is less than nominal, and leakage is extremely low. These indicate excellent performance and reliability. As an example of exceptional reliability in the field, Fig. 2 shows an electrolytic capacitor which was manufactured in August, 1937. It is rated 12 mfd, 450 v dc. It was put in service in a new radio in September, 1937. It was the first filter capacitor in a capacitor input-filter circuit.

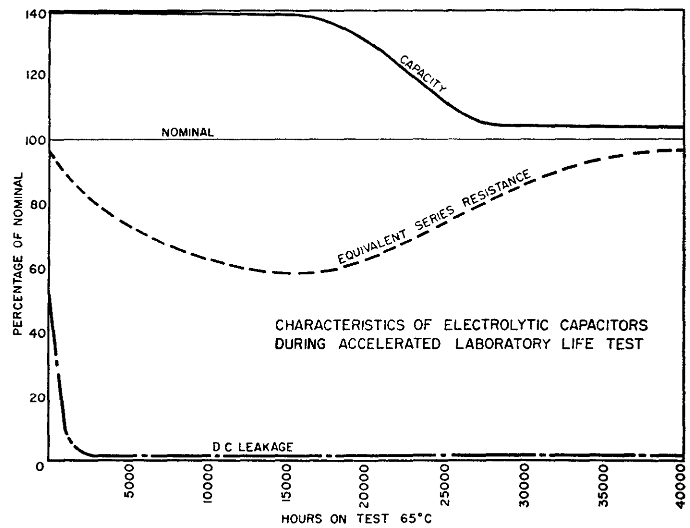


Fig. 1—Characteristics of electrolytic capacitors during accelerated laboratory life test.

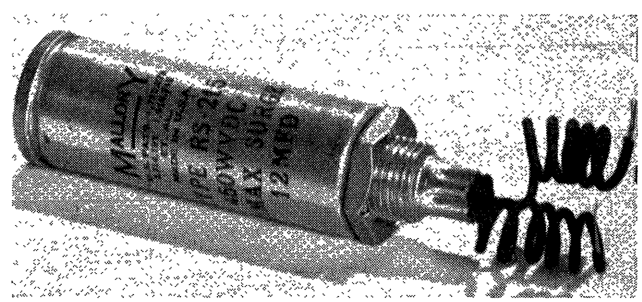


Fig. 2—Electrolytic capacitor after 16 years in service now tests as good as new.

Applied dc voltage was slightly less than rated. It was removed just a month ago even though it still was operating satisfactorily. The radio with this capacitor received a great deal of use for four and one-half years, then was put in storage for four years. When it was removed from storage, it was plugged into an electrical outlet, turned on, and it started to play within 30 seconds. Since then the radio, and the capacitor, have

† P. R. Mallory & Co., Inc., Indianapolis, Ind.

received approximately the same amount of use per month as before the storage period. This is 12 years of service from this capacitor in addition to four years of idle shelf life for a total of 16 years life. The idle period during storage represents a particularly adverse operating condition.

After all this time, the capacitor meets all the test requirements of a new capacitor. Now the capacity is more than nominal, being 12.7 mfd, the leakage is very low, 0.3 ma at 450 v dc, and the equivalent series resistance is very low, 5.7 ohms. Its power factor is 5.4 per cent, which is low even for a new aluminum electrolytic capacitor. There are many other known cases of reliable service obtained from electrolytic capacitors; however, this one is of particular interest because all the details of its life are known. It is of interest to know that this capacitor was picked at random from production, so it received no special care in manufacture other than all capacitors received which were made at that time 16 years ago. One important fact is worth noting here. This capacitor would not have lasted five years if impurities had been present which cause internal corrosion. The reliable capacitor manufacturer keeps these impurities out of his capacitors, so that this long reliable life may be attained.

Once it is known that electrolytic capacitors can give reliable operation, the problem is not necessarily solved. No doubt all of you have a great deal of knowledge of the steps necessary to assemble many components into a completed device so that both components and completed device operate with maximum reliability. Many articles have been published along these lines, so it is not necessary to repeat basic knowledge. As you would expect, reliable operation of electrolytic capacitors is the result of careful and correct engineering, design, and manufacture of the component, followed by careful and correct engineering, design, and manufacture of the complete equipment in which the capacitor is used. It then is necessary to handle, transport, store, operate, and maintain the complete equipment (with its components) correctly and carefully.

Aluminum electrolytic capacitors will be reliable in applications falling within their limitations. The fact that electrolytic capacitors basically are large capacity, small volume devices, at times may lead to the choice of an electrolytic capacitor for an application in which it may not function reliably. The usual minimum capacity is 5 mfd, though a few are manufactured with less capacity. The economical and size advantages of electrolytic capacitors diminish as capacities become less than 5 mfd. The maximum capacity is almost limitless as capacitors may be paralleled to obtain any desired value. Capacitors of 10,000 mfd in a single section at low voltage have been made.

Large capacity in small volume is accompanied by considerable variation in capacity among capacitors made in production quantities. To have reliable operation of a completed equipment in which aluminum electrolytic capacitors are used, the applicable circuits, if low-voltage capacitors are used, must operate reliably with any capacitance within the range of nominal rating minus 10 per cent plus 100 per cent. For high-voltage capacitors this range is minus 10 per cent plus 50 per cent. Also, as electrolytic capacitors may have power factors from 2 per cent to 50 per cent, depending upon the capacitor rating, the circuit must operate reliably with capacitors of this power factor or the over-all operation of the device might be affected. Similarly, the circuit must operate reliably with electrolytic capacitor dc leakage currents from a negligible value to 10 ma, again depending upon capacitor rating.

Just as the circuit must operate reliably when the characteristics of the capacitor are considered, so must the capacitor operate reliably when the characteristics of the circuit are considered. In this respect, the electrolytic capacitor is included in the circuit to perform a certain function. As it performs its function, direct current voltage may be applied to it, an alternating ripple potential may be superimposed on the direct current potential, or it may be charged and discharged periodically. The capacitor must operate reliably with one or all of the above conditions present. There is a limit to the various voltages and currents which can be applied to a given capacitor. Most electrolytic capacitors used in electronic circuits are of the polarized type for application of direct-current voltage in one direction only. For this type of capacitor there is a maximum surge voltage rating and a maximum continuous voltage rating. The voltage ratings for an electrolytic capacitor used in a computer where peak reliability is desired should be less than for the same capacitor used in a television set in a private home. Table I shows a comparison of the standard surge and operating voltages of typical aluminum electrolytic capacitors with the recommended "computer" surge and operating voltages for the same capacitors. Note that the operating voltages in the computer rating are 80 per cent of the standard with a maximum of 360 volts, and the surge voltages in the computer ratings are the standard operating voltages with a maximum of 400 volts.

TABLE I

SURGE AND OPERATING VOLTAGES OF ELECTROLYTIC CAPACITORS

Capacitor	Standard Rating		Computer Rating	
	Operating Voltage	Surge Voltage	Operating Voltage	Surge Voltage
A	450	525	360	400
B	300	375	240	300
C	150	200	120	150
D	25	40	20	25

Not only must the voltage applied to an aluminum-electrolytic capacitor be limited, but the maximum internal temperature must not exceed a safe value. On the low side, the temperature should be at least 15 degrees C (59 degrees F). On the high side, the temperature of any part of the electrolytic capacitor never should exceed 60 degrees C (140 degrees F). Ordinarily, the factors which cause a capacitor to heat up are: ambient temperature, dc leakage current, ac ripple current, charging current, and discharge current.

These limiting characteristics are chosen with the thought that the computer manufacturer wishes to build a device which can be used for many years without electrolytic capacitor failures. Take a group of one hundred electrolytic capacitors which pass all standard production tests and are reasonably uniform. Record all data and put the capacitors in one hundred finished devices which are operated under identical conditions until all electrolytic capacitors fail. They will not all fail at once, but in some particular order. Looking at the data at this point seldom will show any correlation between initial characteristics of the capacitors and the order of failure. If we could apply any satisfactory test to those electrolytic capacitors now accepted as good units on final test and be sure to pick out the one capacitor per hundred which would be the first to fail in actual service, the limiting voltages and temperatures listed above could be increased. Until such a test is devised, the operating conditions must be such that the weakest capacitor of an accepted group gives the desired length of reliable service.

The computer manufacturer who has determined that electrolytic capacitors will give reliable service in certain parts of his device must transmit appropriate information to a reliable electrolytic capacitor manufacturer in order to get the capacitors he desires. In addition to the information as to capacity rating and voltage rating, the capacitor manufacturer must know how those ratings were chosen and what degree of reliability is desired in the operation of the completed capacitors. While there is remarkably little difference between the standard electrolytic capacitor used in television sets and the special capacitor used where maximum reliability is desired, the capacitor manufacturer wants to take every step possible to insure that the computer manufacturer receives the most reliable electrolytic capacitor it is possible to manufacture. The design of the capacitor follows the best known method at the time, including metal outer case. The manufacture of the capacitor requires careful attention to correct procedure. As the electrolytic capacitor industry still is on the rise, design and manufacturing are improving constantly. A year from now the designs should be better and the manufactured product should be better too.

The problem of transporting completed electrolytic capacitors to the equipment manufacturer never has been serious. Capacitors are sealed in the process of manufacture, so that standard packing methods are used for domestic shipments. In the event electrolytic capacitors had to be protected against contaminating fumes, or something similar, it would be necessary to pack the capacitors in fume-tight, fume-resistant containers (or water-proof, if desirable).

The assembly of electrolytic capacitors in computers should be similar to the same operation in other equipment. The position for the capacitors should be away from large heat producing sources. The position should be ventilated so that the capacitor heat is taken away. Electrolytic capacitors which have not been on voltage for a long period of time may require that a low voltage be applied and gradually increased until rated voltage is reached. Ordinarily, this problem is not encountered in new equipment production. The capacitors need to be mounted securely and insulated from the chassis, if necessary. Lead connections should be arranged for easy assembly. It is important that polarized electrolytic capacitors be connected into the circuit with the voltage in the correct direction. Following assembly in a computer, the electrolytic capacitor should not be exposed to possible damage from subsequent assembly operations. This includes mechanical, electrical, and heat damage. A capacitor with a damaged seal will not give reliable operation.

If everything has been done correctly to this point, the computer should operate satisfactorily and be ready for use. There should be no problem concerning the electrolytic capacitors. Many of you are more familiar with what happens to components in a computer as it is used than we are, so we shall leave that up to you.

After talking about how reliable electrolytic capacitors can be in computers, there should be no need to talk about maintenance. However, anything made by man requires maintenance, so it must not be overlooked. The same care of a capacitor installed in a computer during assembly must be followed during any maintenance procedure, in the event something else is receiving the service. There is no regular service for the capacitor such as greasing an automobile at definite intervals. If any electrolytic capacitor is not operating satisfactorily, it probably will have to be replaced. If lead connections were not good, these might be corrected. If a capacitor has to be replaced, it probably is quite old. This means that if spare parts are on hand for the computer, the electrolytic capacitors probably need to have low voltage applied and gradually increased until rated voltage is reached. This should be done without heating up the capacitor appreciably. The used capacitor may be removed and the unused one mounted and connected in its place. This operation must be done with care so that the connections are made to the proper polarity and the capacitor is not damaged in the process.

It is possible to have spare electrolytic capacitors stored with a computer, but to be unable to use them when needed because of improper storage conditions. Capacitors should be stored in a clean, cool, dry location. Voltage should be applied to each capacitor once each year as described above, starting with a low voltage. Under these conditions, spare electrolytic capacitors should be ready for use at any time.

Although there are several limitations to the application of electrolytic capacitors to electronic circuits, there are definite applications in which this capacitor is the correct choice. In those cases, the electrolytic capacitor is superior to any other type known. For any application requiring performance beyond the limitations of electrolytic capacitors, the attempt to use electrolytic capacitors may bring grief. Certain knowledge of electrolytic capacitor characteristics is necessary in applying them to electronic circuits in order to get maximum reliability. Proper knowledge of electrolytic capacitors is necessary to take care of them properly. To obtain maximum reliability from electrolytic capacitors, much more conservative ratings should be employed than is standard in the television field.

There is a practical goal of reliability which human beings can reach, but not pass. This group is vitally interested in reaching that goal. Therefore, none of you would want to use electrolytic capacitors where another type of capacitor or a difference in circuit would give better total machine reliability. Electrolytic capacitors can be manufactured with the minimum human error. They can be applied to electronic circuits to give maximum life and reliability. Electrolytic capacitors will give longer reliable life in computers than some of the other components. In the proper application, electrolytic capacitors should last at least ten years. Your telephone probably is the best general example of the practical goal of reliability. How often have you picked up your telephone and found it out of order? You have done this seldom, if ever. Electrolytic capacitors can be used in computers and this same high degree of reliability attained. This is a joint problem between computer manufacturer and electrolytic capacitor manufacturer. With a co-operative spirit and effort, this problem can be solved to the mutual satisfaction of all.

Discussion

D. J. Crawford (International Business Machines Corp.): When large total filter capacitance is needed what is the optimum size of capacitors that should be used for maximum long term reliability? For example, 5 farads total at 150v working voltage, negligible ripple current?

Mr. VanBuskirk: This negligible ripple current is the key on this particular question. You can go as high as a 2 inch diameter container and a maximum length about 4 to 4½ inches long. We do make 3-inch diameter capacitors but I don't like to recommend them because they get tight on winding. They're used on applications where they're doing a good job, but for computer reliability I like to say 2-inch maximum. If there is anything that will produce heat such as high ripple current then the diameter goes down to 1⅜ inches.

Mr. Crawford: How much should telephone quality and telephone spec. type of capacitors be voltage derated for maximum long-term life and reliability?

Mr. VanBuskirk: That depends upon whether the design of this particular capacitor is derated from the standard. If it is, then you should go back to the radio and TV standard and derate as we showed you on the chart. That will do the job.

C. T. Schaedel, Jr. (Consolidated Vultee Aircraft Corp., Ft. Worth Div.): Please comment on amount of inverse current per microfarad which will not destroy an electrolytic capacitor.

Mr. VanBuskirk: This is basically ripple current, I believe. If not, I'd like to see you later. If it is ripple current, we go by charts and I have some of those with me.

Mr. Schaedel: How about inverse voltage; that is, positive volts applied to negative terminal?

Mr. VanBuskirk: On a polarized capacitor that will destroy it. It will reduce the capacity, probably over-heat it and drive out all the electrolyte.

J. C. La Pointe (Department of Defense): Is the capacity of an electrolytic capacitor a function of the applied voltage? If so, how much?

Mr. VanBuskirk: Yes, it is. You're getting into design here. If you're interested in knowing about it, I'll talk to you after this session. It is something which I don't believe we should take up with the whole group.

A. R. Garfinkel (Franklin Institute): It is my impression that electrolytic insulators tend to deform as the applied voltage is reduced. If this is true, what is the advantage of derating for computer application?

Mr. VanBuskirk: The answer here is the deforming that occurs due to the reduction of voltage is not 100 per cent. The worst condition of reduced voltage and deforming is on shelf life, where no voltage is applied, and in those cases the film is still repairable to the original value, and therefore, the derating that we show here does have a good effect.

J. E. Palmer (RCA Victor): How does the ripple frequency of the applied voltage affect the reliability of electrolytic capacitors?

Mr. VanBuskirk: Here, we're getting down into personal opinion but I'm pretty sure it's right. The detrimental effect of ripple on electrolytic capacitors is a heating effect and if you have two ripple currents of different frequencies that give the same temperature rise within the capacitor I believe you'll find that there is no difference in the effect on the life of the capacitors. That is, regardless of ripple frequencies, the amount of heat you get is the factor you want to worry about, not what is the frequency.

B. B. Paine (Massachusetts Institute of Technology): Will periodic parameter measurements on operating electrolytic capacitors predict catastrophic failure? What schedule for the test would you recommend?

Mr. VanBuskirk: On this, sometimes yes, and sometimes no. You can make parameter tests, and you noticed on the first chart that there were variations in all three characteristics shown—capacity, equivalent series resistance, and leakage current. Now when the capacity gets too low, you can see the trend if you draw a curve and keep making these parameter tests, and tell when the capacity is going to go too low to do the job you want the capacitor to do. Likewise, you can see when the resistance is traveling towards the point where it's too high to operate properly. On leakage current, when it turns around and starts to go up the condenser is gone So you can make parameter tests for certain things, but at the same time you may make these parameter tests and feel like you have thousands and thousands of hours of good life left in the capacitor and it will go in two hours because it shorts out. There's no parameter test we have been able to make to determine that one. Lots of people have tried. We've tried. Some people have thought they've had the answer, but to the best of my knowledge there is no one that has been proved out with sufficient field experience.

H. Wright (National Broadcasting System): Is it correct to say that electrolytic capacitors in series can be double rated in voltage?

Mr. VanBuskirk: Yes, we make capacitors that are two sections in a single container with two leads coming out on which we do use exactly double rating. For computer use even though there is a slight derating there should be a little bit of allowance made for the two capacitors in series. That is if you take two 300-volt capacitors in series, in computer rating they should be operated at a maximum say of 550 volts.

Mr. Wright: Can it be said that in such operation there are compensating factors tending to greater reliability and protection against surges?

Mr. VanBuskirk: Yes, you have to keep your maximum voltage on a capacitor below a certain level to be sure that you do not get a sparking or breakdown. By having two capacitors in series you can get good reliable operation with a voltage that would break down any presently made single section capacitor. On the other hand, electrolytic capacitors tend to be self-balancing when put in series but there is a limitation to how far you can go. In other words you can't build them up for 10,000-volt operation.

Mr. Wright: Can the change in capacitance following protracted operation at subnormal voltage be predicted? Can you give a rule of thumb formula?

Mr. VanBuskirk: It can't too well be predicted because from all the tests we've made it doesn't follow a set pattern. Again as stated on another question, the worst condition is no voltage. There, if you take a condenser on shelf life—we had one once that was nine years old—the capacity on that unit also was a very few per cent different from the initial value after nine years with no voltage. The capacitance change with protracted operation at sub-normal voltage or with no voltage is not a high enough percentage that you would want to say it's eventually going to an infinite point. Fred Keller of the Aluminum Company of America says there is a definite increase of capacity with operation at sub-normal voltage. I'm giving my personal opinion in the matter. If the capacitor is truly operated; that is, if there is any voltage on it at all within reasonable limits you are not going to be able to detect too big a change in the capacity after this operation at subnormal voltage compared to what it would be operated on full voltage.

E. Seif (Burroughs Adding Machine Corp.): How do you define failure of an electrolytic capacitor?

Mr. VanBuskirk: That's when the capacitor fails to perform the function it was put into the circuit to perform. It can do this by various means, for example, it can short-circuit, it can become low capacity, high resistance, or high leakage.

Mr. Seif: Do you consider as a failure the temporary damage to the electrolytic which is self-healed?

Mr. VanBuskirk: At this point, he is referring to the sparking in the electrolyte. It could raise a lot of trouble in computer operations, but the voltages we tried to recommend here will not give sparking in the electrolyte unless there is some additional failure in the unit, that is, a hot spot developing that lowers the sparking potential of the electrolyte. Any temporary damage which is self-healed you want to eliminate because it does give a temporary surge current, and the type of failures that are self-healed of course are a sparking where the oxide film breaks down or deteriorates and then is reformed by the surge current. Then, of course, in low temperature operation electrolytic capacitors may be cooled down to the point where there is no measurable capacity left. This is a temporary condition and the capacity is as good as new when it is brought back to normal temperature.

C. W. Watt (Massachusetts Institute of Technology): Are MIL-specification electrolytic capacitors good enough for computers? Are non-MIL-specification tubular electrolytics good enough for computers?

Mr. VanBuskirk: With the proper choice of the rating for a particular application it is possible to get excellent reliability from either of these types of capacitors in computers. The same care in the choice of the rating must be made for these capacitors as for any other. Naturally, if the application should not be serviced by electrolytic capacitors, then neither of these types will operate satisfactorily.

There is one drawback to the use of MIL-specification electrolytic capacitors in computers. These capacitors are made the same way that capacitors were made when qualification approval was obtained on the capacitors. There is a reluctance on the part of manufacturers to be changing any of the designs of MIL-specification electrolytic capacitors, as approval from the Armed Services must be obtained prior to inclusion of a change in design. As a result, MIL-specification electrolytics are frozen in design and improvements developed subsequent to such freezing are not included in the MIL-specification capacitors. For this reason a better capacitor usually can be made than the MIL-specification type, as the electrolytic capacitor industry continues to progress.

Generally, MIL-specification electrolytic capacitors should be more reliable than the standard highly competitive designs. Then, in turn, up-to-the-minute designs for computer use with maximum reliability will be somewhat better than the MIL-specification type.

Resistor Reliability—Whose Responsibility? Some Case Histories

JESSE MARSTEN†

Introduction

COMPONENT RELIABILITY is a subject that has had the attention, in whole or in part, of almost every electronics symposium or conference the last few years. The failure rate of equipment, claimed to be caused by component failure, warranted this attention. As a result the cry has been for better and better components. This is good. Components should be and are being constantly improved, just as computers and other electronic gear should be constantly improved.

† International Resistance Co., Philadelphia, Pa.

However, this cry for better and better components has been based on the assumption that the component is the villain of the piece. If the component fails or is unreliable, it is the fault of the component. In discussions on this subject one hears that the component cannot stand high temperature or low temperature, it cannot handle overloads, it is unstable, it changes too much for one reason or another, it opens circuits, it breaks down. And so on and on. And the usual cliches are trotted out, as typified by—"A chain is no stronger than its weakest link," and "The dependability of each component determines the utility of the assembly." In brief, the burden of reliability of electronic gear is thrown on the component.

Now, there is no denying that components, like human beings and even computers, are imperfect. They have their weaknesses and inadequacies. And let it be said at the outset that the component manufacturers recognize the need for improvement and are doing everything possible to eliminate these inadequacies. However, there is another side to this coin of component reliability, and it is the object of this discussion to show this other side. Real life, thumbnail sketches are cited, illustrating examples of component failure and unreliability which have nothing to do with the component but everything to do with its abuse and misuse. The symptom of unreliability is in the component, but the real cause is elsewhere. From this analysis some suggestions follow, which it is hoped will prove constructive in obtaining more reliability from existing components.

Case 1

A $\frac{1}{2}$-watt wire-wound resistor molded in bakelite, resistance 120 ohms, used in a sound system circuit, failed occasionally by explosion, like a firecracker. Normal power dissipated in the resistor was appreciably below the rating. The trouble was ascribed to defective resistors. Careful investigation disclosed that momentary line faults resulted in line voltage of 110 volts appearing across the resistor, producing 100 watts in the resistor, 200 times its rating. With the necessary protective measures, no failures have occurred. Now, it may be said that resistors should be capable of handling overloads. This is true, but hardly 200 times. This is a case where the designer should have anticipated possible faults which would affect reliability and guard against these faults in advance.

Case 2

High-stability deposited carbon resistors were used in an amplifier. The prototype equipment produced by the development company was satisfactory. The ultimate manufacturer found his amplifiers did not meet specifications and eventually traced the trouble to the deposited carbon resistors which were claimed to be unstable. They were considerably outside the original 1 per cent tolerance.

Investigation disclosed that the manufacturer sprayed the resistors with a fungicidal compound which attacked the insulating coating of the resistor and eventually the resistance film, causing high resistance changes. The prototypes were not treated this way. Too often the user assumes the component should be capable of withstanding any treatment he applies to it. Not infrequently the user processes components without consideration or knowledge of the effect of such processing on the component, and without consulting the component manufacturer. Many such failures can readily be prevented by prior consultation with the component manufacturer who can frequently guide the user to the correct materials to use, if necessary.

Case 3

Miniature $\frac{1}{4}$-watt composition resistors with a very thin bakelite molding were found to be mechanically unreliable, which would eventually lead to electrical instability. These resistors, oddly enough, were assembled in the equipment in such a manner that to solder dip the leads it was necessary to immerse the entire resistor, body and leads, in a molten solder bath at 520 degrees F. Although the resistance value did not change too seriously (this change was not the main trouble) the bakelite body showed signs of cracking. Needless to say, resistors of this class are not designed or intended to meet this kind of service.

Here again there seems to be an underlying assumption that resistors or components in general should be able to withstand any kind of treatment to which the user wishes to subject them.

Case 4

Deposited carbon resistors—150 of them—assembled in a computer were claimed to be unreliable in that about 70 per cent of them departed anywhere from 1 per cent to 10 per cent from the original 1 per cent tolerance. The 150 resistors were to be removed and replaced by another make. Here was a situation in which the end equipment had not yet failed but failure was anticipated because of the above assigned reason. Twelve tagged resistors were produced showing deviations of 1 per cent to 10 per cent. A check measurement on a Wheatstone Bridge showed all to be within 1 per cent. The 150 resistors in the computer were measured and only one was outside the 1 per cent tolerance—and the deviation was nominal. Here is a case of guilt by assumption and assumed unreliability, not based on fact or actual performance, but based on inaccurate measurements. The reason for the inaccurate measurement was never determined or revealed.

Case 5

A 2-watt composition resistor, used far below its rating, was giving serious field trouble in that its value changed excessively and the bakelite insulation showed signs of discoloration, pointing to excessively high temperature. This resistor had proved satisfactory in engineering development models, and in the initial engineering prototype samples. Yet all production equipment showed this field trouble. After the usual trouble shooting, a comparison of prototype and production models disclosed an apparently minor difference. In the prototype, the 2-watt resistor was assembled at some distance from a 10-watt wire-wound resistor operating at quite a high temperature. For reasons of of wiring simplification, the factory assembled the two resistors cheek to cheek in production. The result of the direct heat transfer from the power resistor to the composition resistor, the temperature rise in the composition resistor itself, and the ambient temperature, was an operating temperature far in excess of that for which the composition resistor was designed. The correction was obvious. There was nothing wrong with the resistor, there was everything wrong with its mounting. No consideration was given by the manufacturer to the effect change in assembly might have on the resistor.

Case 6

A 1-watt deposited carbon resistor specified at 1 per cent tolerance departed sufficiently from its tolerance limits in the circuit application to impair performance. In operation the resistor changed about 3 per cent to 4 per cent which was claimed to be objectionable. Operating at full load this change is not bad for a deposited carbon resistor, regardless of its initial tolerance. This class of resistor has a relatively high temperature coefficient—about 200 to 600 parts per million depending on resistance value. At full load the temperature rise is about 80 degrees C. Depending upon resistance value this resistor could change as much as 4 per cent because of temperature rise alone. And if the ambient temperature were appreciably higher than 25 degrees C. the change would be greater.

The difficulty here is that many engineers confuse or equate close tolerance with stability. One has nothing to do with the other. It is possible for a 20 per cent resistor to be far more stable than a 1 per cent resistor. In this particular instance, if temperature stability was as important as claimed, a resistor of lower-temperature coefficient was clearly indicated, or a larger deposited carbon resistor of the same type used below its rating should have been used. Unfortunately, analysis of all factors involved in the use of components is frequently not made, and the path of least resistance is to blame the component for the failure of the user engineer.

Case 7

Trouble was experienced with a special type of resistor presumed to be a high-stability type. This time the trouble was real; there were wide deviations from the nominal tolerance, which were definitely traceable to the inadequacy of the resistor for the application. However, the resistor mounting was so tightly designed that it was almost impossible, without a major operation, to mount the desired replacement resistor, although it was only 1/32 inch longer than the original.

Innumerable other instances may be cited. A few odd cases of trouble may be mentioned in passing, such as the case of the laboratory which in the year A.D. 1952 was still using composition resistors of 1937 vintage; and the case of the guided missile model, in which resistors were used which had been soldered in and out of breadboard models innumerable times; and the case of chronic fires in projection rooms of motion picture theaters traced to prettily cabled wiring with flammable insulation, carefully dressed and draped around power wire-wound resistors operating at nearly 250 degrees C.

These, then, are just a few samples of what happens to resistors—and undoubtedly other components—which cause equipment failure and unreliability. We mount and wire components in almost complete disregard of the possible consequences; we process them in ways which may be destructive, without inquiring as to the effect of such processing on the component; we use minimum space in designing resistor mounts thereby making it difficult, if not impossible, to substitute more reliable parts in the event of trouble; we ignore available technical data and use close tolerance resistors in the hope that they will solve our stability problems; we do not analyze circuitry to anticipate possible troubles and provide protective measures in advance. In brief, we assemble the resistor or component in the circuit and expect it to carry the reliability burden.

These instances of abuse, misuse, and misapplication of components sometimes raise a certain skepticism as to whether reliability is as important as we are led to believe, because the treatment to which these components are sometimes subjected is the direct opposite to that which reliability would dictate. They illustrate quite conclusively that component reliability is not the exclusive responsibility of the component manufacturer, but is equally, if not more so, the responsibility of the component user. What he does to, and with, the component is often the principal determinant of the reliability of his equipment. If due regard were paid by the user to the properties and limitations of components, the incidence of failures would be substantially reduced.

It is interesting to note that this factor has been largely ignored in most of the discussions of component reliability. The component has been considered a thing apart, unrelated to its environment. It is good or bad, depending upon what is in it. This is only part of the story. This attitude ignores a most important factor in the component environment, namely, the human element around it—the research, development, design, and factory engineers. What they do can make a reliable component appear unreliable—witness the cases here cited. Or their actions can make an apparently less reliable component be very reliable by proper use of the component and proper original circuitry design.

This latter is beautifully illustrated in a paper entitled "Rudiments of Good Circuit Design," by N. H. Taylor, of The Massachusetts Institute of Technology, given at the April 1953 Symposium on Component Reliability at Pasadena, Calif. Two designs of flip-flop circuits are described. One, conventional, employing 1 per cent resistors; the other, original, employing 5 per cent resistors and a few more components. The latter had a much greater reliability even though it utilized components which varied over a wider range of values. The important point here is that Mr. Taylor set out to design a more reliable system avoiding the use of close tolerance resistors and using more available and, perhaps, less stable components. All of which leads to the main point of this discussion, namely, objectives.

There are no perfect components, so we must learn to live with what we have. But the choice is great. Also, we have to know what we want when we make this choice. For example, there are at least five different types of ½-watt resistors: composition, deposited carbon, boron carbon, metal film, molded wire, and precision wire. Every one of these fills an important need and has its place in the electronic scheme of things. They vary widely in cost, size, range of values, and characteristics. But none of these resistors has everything. The choice must therefore depend on the characteristics of maximum importance. The user-engineer must know his requirements. And this presupposes that he knows the operating conditions in his equipment (which unfortunately is not always the case). With this in mind a choice is made. There is no point in asking for everything, smallest size, low cost, low-temperature coefficient, ability to withstand high temperature, minimum change with time and humidity, etc. There is no such thing. Somewhere a compromise must be made. If the choice is properly made with all the facts at hand, the results should be good. If the results are not good, the choice may not have been correct, or that is the best that can be done in the present state of the art, or the objective of reliability was not uppermost in the mind of the user.

The case of the 1 per cent deposited carbon resistor (Case 2), which changed about 4 per cent under load, illustrates this point fully. If temperature stability were the prime requirement, and such a change were intolerable, the wrong resistor was chosen. This class of resistor is reliable and stable, but it cannot meet this particular requirement. There are resistors that can, but a price would have to be paid either in size or cost or both. If this requirement was paramount in influencing reliability of the equipment, then the price should have been paid. If other factors were predominant, then the resistor should not be condemned as being unreliable for it was doing the job it was designed to do.

A new piece of equipment currently calls for resistors of the order of magnitude of 50,000 megohms. Apart from certain electrical requirements, there is a size limitation. It must be preferably as small as a ½-watt composition resistor, perhaps as large as a 1-watt composition resistor. The size and value practically dictate a composition-type resistor, either solid or film, with the reliability inherent in this class of resistor. If for any reason the equipment did not perform properly because of the expected changes in resistance, it would not be correct to say this resistor was unreliable. It is perfectly reliable within its limitations. Even though reliability was of paramount importance, there is no other choice. We must live with the best compromise available.

A final point—objectives in equipment development and design. What are the objectives usually set forth in a new equipment project? Performance, always—equipment must meet specified performance standards. Most often, space-miniaturization has almost become a fetish. Cost, almost always. These are basic objectives and are always spelled out. No doubt there is some thought of reliability. But how often is reliability spelled out as a prime objective, if not the prime objective, in equipment development? How often is the specification established that above all else the equipment shall not fail, even if it means sacrificing some performance, space and cost? It is the belief of many of us in the component industry that if reliability were established as a basic requirement, equipment engineers would approach this problem differently. They would exercise more care in the assembly of components; they would study the limitations of components more carefully and use them in ways which would not tax them unduly; they would devise systems and circuits, perhaps unorthodox, which would permit use of components with wider tolerance and variations; they would use more and larger components, if necessary, to insure reliability. The burden of responsibility for reliability would then be more equitably divided, and the incidence of unreliability considerably reduced.

Discussion

C. T. Schaedel Jr. (Convair, Fort Worth): In reference to your figure of 200 to 600 ppm, per degree C for deposited carbon resistance, is this positive or negative?

Mr. Marsten: Negative.

H. Rosenberg (Burroughs Corp.): Applying a duty factor which causes an average dissipation within rating, what is the peak dissipation allowable with standard half-watt, one watt and two watts, 5 per cent and 10 per cent resistance?

Mr. Marsten: I presume reference is to composition resistors, and I also suppose this question refers to pulsing or perhaps surge. I can't answer the question. There is a big gap really in the information available on peak power dissipation and it is a gap we hope we're going to fill one of these days.

Reliability and Its Relation to Suitability and Predictability

E. B. FERRELL†

RELIABILITY, like a great many other words, means different things to different people. Let me illustrate with a purely imaginary example. Suppose we have a small vacuum tube with amplification, mutual conductance, and plate impedance all of useful magnitudes. Tubes of this type have been made in large quantity. Their characteristics, when measured at the factory have very good uniformity—all are within ±1 per cent of their nominal value. Every tube that has been examined has kept its characteristics within these narrow limits throughout its entire life, and a large fraction of the tubes made have been thus examined. These characteristics are entirely independent of such things as ambient temperature and mechanical shock.

Now we have, or did have, two potential customers for this tube. Both had specified the same mutual conductance, the same plate impedance, and so on. After extensive trials one of the customers adopted our tube, uses it now and is quite happy with it. The other very quickly turned it down.

Our satisfied customer is working on a classified project—a bomb fuze, I believe. The other builds home television sets.

Oh, yes! I forgot to mention that this tube, every one we have tried, will operate for just 10 minutes, and then blows up—explodes—scatters little bits of glass among the equipment it's mounted with.

Is this tube reliable? Well, you can certainly depend on it. It acts just the same way every time.

I looked up the word reliable in the dictionary. "A thing is reliable when one can count on it not to fail in doing what it is expected to do." I don't like that definition. I don't think it obeys a fundamental engineering concept that the properties of a device must depend on the device and not on the person who is talking about it.

But still, that is the way we use the word, or almost the way we use it. One customer says my tube is reliable the other says it is not. The television set man has come to expect the tube to blow up and it does. Therefore, by the dictionary it is reliable. But he had hoped the tube would last longer and it doesn't. Therefore, he calls it unreliable. Our definition should read: "Count on it to do what we hope it will." Even amended, I don't like the definition. Perhaps we should break this word reliable up into two parts: if a device is to be reliable it must be predictable and it must be suitable.

The simple phrase "The tube is predictable." makes sense. True, to prove such a statement you must specify certain tests, I must describe the results of these tests in advance, the test must be performed, and the results must agree with the "predictions." But to say the tube is predictable means that this statement and proof can be given, whatever the test conditions.

† Bell Telephone Labs, Murray Hill, N. J.

The phrase "The tube is suitable." doesn't make sense until you complete the statement "suitable for what?"

In order to be reliable, a device must be both suitable and predictable. Since suitability depends on the job, we should never say a device is reliable, and stop. We should say a device is reliable in a certain job.

And since reliability depends on the job, we should not ask the producer to produce a reliable device. We should ask him to produce a predictable device, with certain characteristics. Reliability depends on needs.

Now "suitability for a job" implies a certain performance in the future. Therefore, if a device is to be suitable for a certain job it must be predictable with respect to certain properties. If a device is suitable it must be predictable. But predictability does not guarantee suitability.

I think that in the final analysis the crux of this reliability business is the predictability. I believe most of our discussions, in some cases arguments, about reliability are really discussions of predictability.

About prediction I would like to quote from Edwards Deming, recently retired after long service as statistician in the Department of Agriculture, the Census Bureau, and the Bureau of the Budget:[1]

"An inference, if it is to have scientific value, must constitute a prediction concerning future data. If the inference is to be made purely with the help of the distribution theories of statistics, the experiments that constitute the evidence for the inference must arise from a state of statistical control; until that state is reached, there is no universe, normal or otherwise, and the statistician's calculations by themselves are an illusion if not a delusion. The fact is that when distribution theory is not applicable for lack of control, any inference, statistical or otherwise, is little better than conjecture. The state of statistical control is, therefore, the goal of all experimentation."

So let's talk about predictability. Here is another example—not fictitious and absurd, this time, but very real. At the beginning of 1942 we were working on a small relay. This relay appeared to have a number of desirable characteristics: speed, sensitivity, load capacity, and the like.[2] We were making these relays in

[1] From the editor's foreword to "Statistical Method from the Viewpoint of Quality Control," published by the Graduate School, U. S. Department of Agriculture.

[2] This example and its statistical aspects are discussed more fully in "Statistical methods in the development of apparatus life quality," *Trans. AIEE*, vol. 64; November 2, 1945. The charts shown here are from that paper.

small batches—batches of 20, I believe. By studying the variations within the batches, and by comparing different batches, we had established a set of standards for our process. These were not standards anybody had asked us to meet. They were what the process itself had told us it was capable of doing.

Oh, we didn't always produce relays that met these standards! But we found that whenever we failed there was some simple reason for it. When some failure in the process made one switch in a batch bad, there were usually several bad ones in the batch. And when a whole batch, individually and collectively, came up to the standards that the process had set for itself, we had faith in those switches. We believed that they were predictable.

Actually, this criterion for control, for statistical stability of the process, was the simple quality-control chart. This consists of taking the relays in small subgroups of 4 or 5, in the order they are made, measuring, say, their operating currents, plotting the difference between the largest and smallest currents in the group on one chart, and plotting the average of the largest and smallest on another chart. See Fig. 1. Simply-computed limit lines are drawn on each chart. So long as the plotted points stay inside the limits we have faith in the consistency and predictability of our switches.

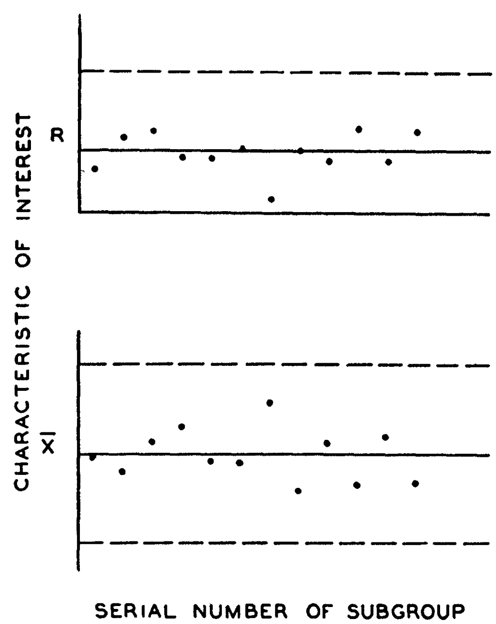


Fig. 1—A control chart.

Whenever a point goes outside the limits, the whole batch is suspect.

This, then, was a method of selecting relays that we thought were predictable. Having gotten to this stage, we made some life tests. And I don't mean death tests. I mean living, aging tests.

We put a few of these relays on test under certain loads. They were operated 60 times per second. At certain intervals they were removed from the work circuit and measured. We measured them when we started the test. That is, we measured their operate, or pick-up, current; and we measured their release, or drop-out, current. We measured these again after they had run an hour. Then after 3 more hours. Then we made measurements twice in an 8-hour working day. Soon we let them run overnight—two measurements in 24 hours.

TABLE I

LIFE TEST INSPECTION SCHEDULE

Number of Inspection	Age in Hours	Added Age	Number of Days	Number of Groups (2)
1*	0			
2†	0	0		1
3	1	1		
4	2	1	1	2
5	5	3		
6	8	3	2	3
7	11	3		
8	14	3	3	4
13	33	7		
14	40	7	6	7
15	47	7		
16	54	7	7	8
21	89	7		
22	103	14	10	11
23	110	7		
24	124	14	11	12
29	188	22	14	
30	210	22	15	15
31	232	22	16	
32	254	22	17	16

* At completion of manufacture.
† At start of life test.

After this we measured them every time the age had increased by about 10 per cent. When they were 10 months old we were measuring them only once a month. This schedule is shown in Table I.

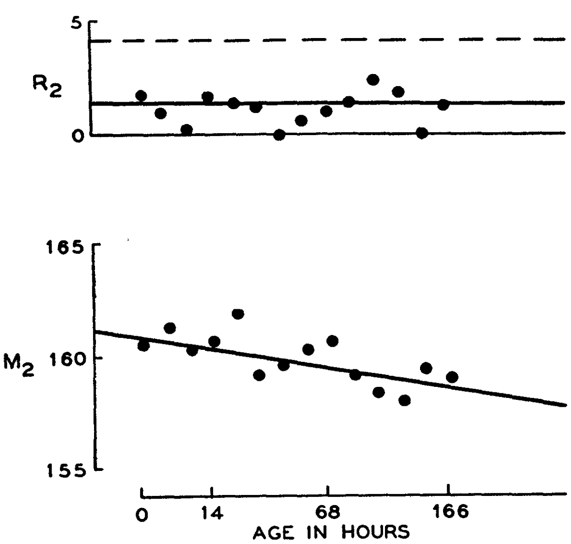


Fig. 2—Operate current—relay 740.

Fig. 2 shows the results of the first 28 measurements of operate current during the life test of one of the relays. Each point on the upper chart is the difference between two consecutive inspections. We call it a range of two, R_2. Each point on the lower chart is the average of two consecutive inspections, the mean of two, M_2.

The points are spaced uniformly along the abscissa. Except near the beginning this is roughly a logarithmic scale of time. The ordinate is ampere-turns to operate. On the range chart we have the usual control chart limits. The points are well inside these limits. This indicates that the short-term variations between measurements is self-consistent and hence predictable.

On the lower chart, the chart of averages, we have not drawn the limits. If we had, points would be outside. These outages indicate that in the long run there is some definite reason for disagreement among the measurements. Somewhat hopefully, we have drawn a sloping line down through these points. Maybe it indicates an aging trend.

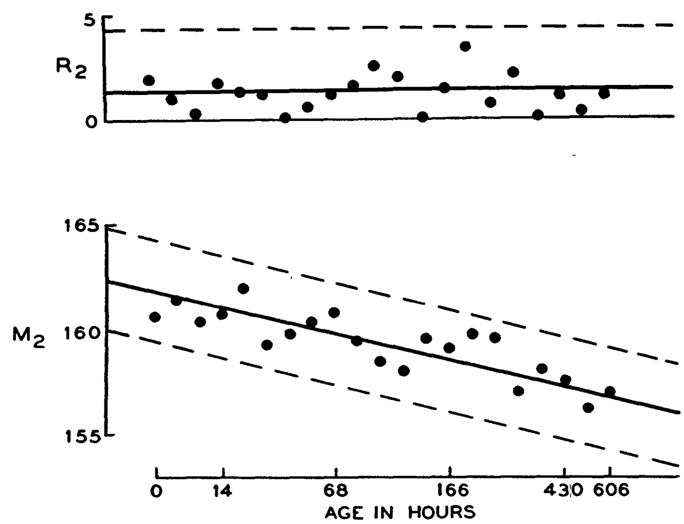


Fig. 3—Operate current—relay 740.

In Fig. 3 this record has been carried through 606 hours, as compared to 166 hours for the previous figure. The trend line has the same slope. We have enough faith in it now that we have drawn limit lines parallel to the trend. Actually the slope of this trend is the average slope for 10 relays. They all stayed within limits placed with respect to this average trend. The ranges, or short term variations are still within control.

Here we have found one reason for differences between measurements made a long time apart. This is simple aging. We represent this by the trend line. Otherwise the measurements are in agreement. We think we are demonstrating predictability.

How long do we keep this up?

Fig. 4 shows the results at 5,000 hours. The top curve is operate current, which was shown in Figs. 2 and 3. The slope of this line, an average of 10 relays, is the same as that determined at 166 or 606 hours.

The report for which these figures were prepared was written at 5,000 hours. Actually, the test was continued to 15,000 hours. Between about 12,000 hours and 15,000 hours—this is in the neighborhood of 2.5 billion operations—nearly every relay gave readings outside the control limits. Shortly after this the test was stopped for other reasons. None of these 10 relays ever actually failed. But their characteristics became erratic and unpredictable. We considered that their useful life was over.

This is my example. We had a method for telling whether relays as made were consistent with each other—the simple control chart. We had a method for telling whether the relays aged in a manner consistent with themselves and with each other—a control chart plotted at a slope on a linearizing scale. These control charts are means for testing and displaying statistical consistency—a method of bookkeeping if you please. Statistical consistency, we firmly believe, is synonymous with predictability; and predictability is the first component of reliability.

Was this the perfect experiment I have painted here? Of course not. As the test went along we got occasional points out of control. These were red flags that something was wrong. One time we found a carbonized resistor in the test set. Another time the test set went bad on a day of high humidity.

Actually we started 12 relays on this test. Two of them became erratic at about 40 hours. We found their work circuit had been wired wrong and they were getting abnormal punishment. They actually failed at about 600 hours.

We started other tests after this one. At loads of some 10 times those of these tests we got erratic behavior after about 500 hours.

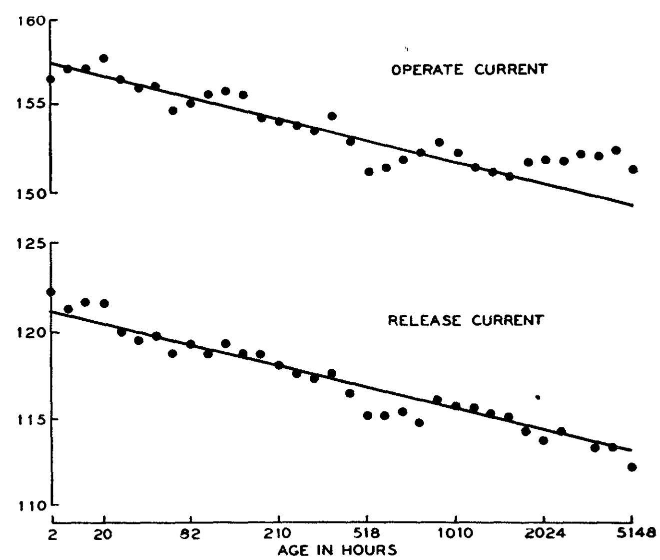


Fig. 4—Operate and release currents—average of ten relays.

In one test of two relays we got an aging trend of a different slope. This pair tested individually and initially as good relays, but came from a batch that was known to contain bad relays—a demonstration of our idea that the relays in a batch must be individually and collectively good to indicate predictability.

Let me close with a plea to the producer of component devices: Make your product first of all predictable. The circuit boys are smart. Within reason they can tailor their circuits to use your product if they can only depend on it to do what you predict it will. But if it is not predictable your product is of little use to anybody.

Discussion

W. K. Halstead (RCA): What can be done to convince manufacturers to supply statistical life or quality control data to the user of the products?

Mr. Ferrell: You're asking a question on human relationships, which is outside of my field. I don't know of any way to persuade manufacturers to give you good quality control data except to convince them that they won't sell you their product unless they do. I do think it would help if the users of the equipment would make logical use of this data. I think in many cases they don't.

Summary of AIEE-IRE-ACM Conference

ALLEN V. ASTIN†

Introduction

TO ATTEMPT to summarize the disclosures of this important three-day conference on information-processing systems is for me a most difficult assignment. Each paper was in itself a summary of extensive investigations, usually covering the co-operative efforts of large teams of engineers and scientists. Furthermore, I was unable personally to hear more than a few of the papers although I was fortunate in having advance copies of most of the others. For the remaining coverage I am indebted for effective reports to the good efforts of Miss Mary Stevens and Sam Alexander of the NBS staff. I am further indebted to these two for the opportunity of discussing highlights of many of the conference papers.

Probability the most significant feature of this conference is the overwhelming evidence of tremendous progress which has been made with information processing systems in ten short years. It is almost exactly ten years ago that construction on the first large-scale electronic digital computing machine, the ENIAC, was initiated. It is only a little more than 3 years since the first full-scale productive operation of a large high-speed automatically-sequenced machine, the SEAC. Now there are approximately 50 of these powerful data-processing systems in use or available for use. Estimates of such an impressive increase in our capacity to make computations and process information would probably have been treated skeptically in 1943 even by those who conceived and planned the ENIAC. For us now in 1953, it is most certainly an occasion for stock-taking. The theme of this conference, reliability and requirements, is essentially directed to such a purpose. Accordingly, the conference papers have been concerned with a critical analysis of how dependable and effective these new tools are and what needs to be done to utilize them and the basic techniques embodied in them more efficiently and more widely.

† National Bureau of Standards, Washington 25, D. C.

In spite of the rapid progress and growth of our information-processing systems, I believe it would be presumptive and noncritical to claim that they have reached maturity. It would probably be more realistic to class the present state or growth as the adolescent stage—major physical growth has been accomplished and there lies immediately ahead intensive training, specialization and refinement in order that this great, somewhat raw potential strength can be fully and effectively utilized.

In analyzing requirements for future development, it is natural to examine present utilization to determine gaps and possible trends. Of the present half-a-hundred major machines, the overwhelming majority are devoted to filling needs of the Federal Government, either directly in government-operated or -owned facilities or indirectly in universities or industrial plants where work on government contracts provides the primary incentive and need for the machine. Although the present situation is reasonable and understandable because of the nature of the incentive for the original development of the machines, it is obvious that if we are to have balanced utilization of their capabilities, the primary areas for new applications are in industry and business. A number of the problems involved in the extension of machine techniques to these areas have been clearly given in Mr. Davis' paper, which was ably presented to us by Mr. Finelli. One of the important requirements for more effective business utilization is greater specialization to meet specific requirements and less emphasis on all-purpose machines. Related to this, Mr. Davis suggested more attention to pilot-line studies and less effort on full automation.

It seems apparent that if machine techniques are to be brought to bear to handle information-processing problems in the many unexploited new areas, there will need to be much more attention to the development of more highly specialized devices, and probably simpler and cheaper devices, to meet specific operating requirements. The various applications call for a wide range in capacity, speed of response, and reliability. To solve economic problems of the input-output type and related logistic problems, problem-solution times of days and perhaps even weeks can be tolerated. To handle effectively the weather prediction problem, solution times of the order of a few hours appear to be necessary. To handle many of the inventory control, accounting and collating operations of business offices, problem-solution times of the order of minutes will be necessary. To handle real-time control systems of the sort described by Mr. Conn, and for the air traffic control problems, as described by Mr. Weihe, problem-solution times of the order of fractions of a second will be necessary. In general, but not always, the requirements for memory capacity are greater with the most slowly solved problems and the requirements for accuracy and reliability are higher with the problems requiring the most rapid solutions. In the air traffic control problem, for instance, computations have to be made of alternate patterns and paths of events expected to occur in the new few minutes or even seconds. On the basis of these computations, irreversible decisions will be made which may affect human life. Thus reliability and speed are of paramount importance.

Mr. Davis' paper pointed up another factor which may become increasingly important in extending the applications of these machines. In general, the capacity of these large, general-purpose tools is so great that it is difficult for a single business firm to use one fully or efficiently. Furthermore, their cost is so high that only the very largest of business organizations can afford the investment. This points to the need for considering seriously the possibility of central information processing facilities where machine time could be leased to relatively small-scale users. If machines of increased mobility could be developed and made available, then machine rental on a weekly or even a daily basis might be practicable. It is, however, probably more realistic to consider for this purpose fixed machines with highly flexible and extensive communication systems.

The requirement for increased ease of communication between numerous widely dispersed sources of raw material or data, and a central information-processing facility is becoming more evident. An important conclusion from Mr. Smagorinsky's paper is that the communication problem between the point where weather data are collected and where they might be centrally processed is fundamental to the ultimate satisfactory prediction of weather by automatic information processing techniques. The collection, sorting, summarizing, transposition and distribution of the weather data by human operators introduces serious time delays and serious sources of error. Somewhat similar communication difficulties occur in the air traffic control problem either for military or civilian application and also in many business accounting and inventory problems where an agency has widely dispersed sources of raw material or points of distribution. There seems to be most certainly a requirement for increased emphasis on the development of widespread communication networks specifically adapted to transmitting information to and from centralized data-processing facilities. Perhaps it is possible that a single communication network could be used for a wide variety of applications on a lease basis. Furthermore, it may be that only through co-operative utilization of such a communication net would it be possible to support it economically.

A number of the papers at this conference were concerned with reports on reliability on the basis of operating experience of the machines now in use. Two types of reliability appear to be involved, one related to the availability of the machine in working order for the solution of problems and the other related to the accuracy of the information processing itself. In general, the reports indicate a high degree of reliability on both scores and there appear to be no outstanding differences between the various machine types. Availability of a particular machine for useful work at least 60 per cent of the time appears now to be commonplace and the present upper limit appears to be somewhere between 80 and 90 per cent. There appears no good means of measuring reliability with respect to accuracy although several methods were proposed. In general, however, this accuracy seems almost unbelievably high. Millions and millions of consecutive computations can be made without error. Furthermore, there have been very interesting developments, as shown by Dr. Mauchly and Dr. Goldstine, in the development of techniques so that what few errors of this kind occur, can be checked or corrected by the machine itself either by built-in checking techniques or by programming techniques.

It is gratifying to observe that the primary difficulties which delayed the utilization of electrostatic memory devices appear to have been largely overcome although there still is substantial room for further development and improvement. It is also gratifying to observe that substantial progress has been made in the development of magnetic core memories. Perhaps such devices are the answer to the problem where essentially perfect accuracy is required.

There appears to be no outstanding differences in the causes of down-time as reported for the various machines. Mechanical failures and tube failures predominate in the lists of causes of trouble. This is, of course, not surprising in devices composed of thousands of vacuum tubes and diodes and containing hundreds of thousands of soldered connections. It is furthermore interesting to observe that the solid-state devices, as represented by diodes, are by no means immune as sources of error and difficulty, although many of the difficulties experienced with germanium diodes appear to be of the type which might be eliminated if they were better packaged and hermetically sealed.

There has been encouraging progress in the developments related to a number of the critical components employed in information-processing systems. In particular the development of better techniques for producing magnetic tapes, as reported by Mr. Wetzel, resulting in less than one error in five reels of tape, is highly significant.

In the earlier phases of machine development, there was an appreciable tendency to carry the development of components only far enough to secure minimum performance characteristics. Marginally acceptable components are probably now the greatest obstacle toward increasing reliability of these machines. As pointed out by Goetz and Geisler, many machines require diodes whose performance characteristics are on the limit of those that can be readily produced by present production techniques. In order that we may make substantial progress in increasing the over-all reliability of information-processing systems, it will probably be necessary to devote more and more effort directly to component development and attempt to obtain components whose performance characteristics are well in excess of marginal limits for machine performance. The work of the RETMA, JETEC Tube Committees, as reported by Mr. Briggs, is certainly a step in this direction and their emphasis on the necessity of attention to the time dependability characteristics of tubes will probably bring about significant improvements in these critical components.

With components generally, more satisfactory results will undoubtedly be obtained if greater attention is given to predictability and consistency of performance, as suggested by Mr. Ferrell, or to more definitive specifications, as suggested by Mr. Van Buskirk, or in taking more care, as suggested by Mr. Marsten, that the actual conditions of use do not introduce unexpected or improper exposure conditions. Essentially the recommendations of these last three speakers have much in common: if sound engineering techniques, coupled with some common sense are used in the selection, specification and use of common components, many difficulties affecting over-all reliability will disappear.

It is not practicable in this summary to refer adequately to each of the excellent papers that was given. Each of them did, however, contribute significantly to the theme—Information Processing Systems—Reliability and Requirements. Certainly acceptable reliability has been established. There is much room for progressive improvement of reliability; but as shown by the papers, those concerned with and able to do something about improved reliability are aware of the problems, and they are not idle. Relatively, the requirements picture is not in such good shape. Much needs to be done before the great potential benefits of these powerful new information-processing systems can be more fully utilized not only in science and engineering where the primary initial benefits have accrued but in factories and production lines and in many business office types of operation. Also there is need, if more balanced participation in the benefits of the use of these new tools is desired, of greater utilization of the techniques by private organizations in their own operations. To meet these requirements, many different type problems must be solved, involving not only specialized engineering development and related research, but economics, education, and co-operative efforts between engineers and mathematicians on the one hand and accountants and businessmen on the other. The solution of such problems can be greatly facilitated by more systematic and effective communication among these groups. For this purpose the Joint Computer Committee of the three professional societies sponsoring this conference serves a most useful function and it is the logical group to take the initiative if we wish to extend the potential benefits of modern information processing techniques to a wider group.

I would like to congratulate the members of the Joint Committee for the great strides it has made in disseminating information about, and creating interest in, electronic data-processing systems. I would also like to congratulate the local committee for arranging a comprehensive program and a stimulating meeting.

The subject matter of this conference is of immeasurable importance in three vital areas—in our national security, in our domestic economy, and in extending the frontiers of knowledge in the natural sciences. The first and last of these have already been accelerated by the impact of this new technology and are depending heavily upon it for future advancement. The second area has yet to experience in any substantial way the benefits that are surely ahead. There are many who believe that the ultimate benefits to be derived in this slower-to-be-developed area will far overshadow those in the areas that are now ahead.

To those of you who are working actively in this exciting field of science and technology and who have accomplished so much in such a short time, I offer my sincere thanks and congratulations. May your future achievements be as creditable as your present ones.

Group Discussion on Diagnostic Checks

J. J. EACHUS†, *Moderator*

A GROUP of about 60 conference members met to discuss diagnostic checks. Many of the conferees contributed to the discussion but in this summary no attempt is made to associate the individual members or their affiliation with their comments.

Discussion opened, logically, with an effort to define the distinction between marginal testing and diagnostic checking. The effort was not entirely successful, it being indicated that diagnosis to find what is wrong is not readily separable from systems testing to find if anything is wrong. It was the general feeling that "diagnostic circuitry" associated with the marginal-checking routine should provide the diagnosis. Purely diagnostic tests may, however, be made without marginal checks by special routine programming.

It was brought out clearly that the greatest interest was in connection with marginal testing for preventive maintenance. Further, such tests are nearly all attempts to forecast the lowering of mutual conductance of tubes below the failure point, and in particular for the twenty-four hour period succeeding the test. No great concern was shown over details of methods to locate "solid" faults, the opinion being that while the method will depend on the particular machine, methods are in each case readily available. A question was raised as to the advisability of keeping a record of the actual limits of various machine sections or circuits under marginal test and using the change or rate of change of these limits to predict failure. One large company is currently trying this method. A representative of one college said that they had used this in exploring for test limits but that for routine operation a fixed-variation test schedule is easier to mechanize and that record keeping is greatly simplified. Fear was also expressed that continual varying of the actual limits may damage the equipment.

It was the general but not unanimous opinion that components other than tubes and diodes are sufficiently reliable that testing of these, other than at long intervals and outside the machine, is not necessary. Divergence of views and of practice was indicated with respect to the details of voltage variation for marginal testing; that is, which voltages should be varied, should the variations be step-wise or continuous, should increments be both positive and negative with respect to the standard, and should more than one voltage be varied at a time. There was a wide difference of opinion on the value and desirability of variation of heater voltage as a testing means. A representative of a tube manufacturer stated that testing with lowered heater voltage is a good way to detect cathode inter-face troubles.

† 2802 Wisconsin Ave., N.W., Washington, D. C.

Concern was expressed over the lack of methods of testing diodes in place. It was suggested that where diodes are used as a self-controlled input switch to circuits such as binary counter stages the marginal tests applied to the tubes may reveal bad diodes. Step-voltage variations in the bias supplied to diode logic circuits was also suggested as a possibility for these circuits. In general, it was agreed that marginal checking of diodes is a difficult problem.

An appeal was made for suggestions which might lead to an easier diagnosis of troubles which occur only under dynamic conditions or in a special sequence of operations. No diagnostic methods were suggested for troubles of this sort arising from variable duty cycle, addition or reflections in delay lines, variation of load on individual circuits or power supplies, other than to apply marginal conditions in order to transform them into consistent failures. In some cases detailed monitoring possibly with an oscilloscope is the only answer. In this connection experience was cited that variation of pulse rate sometimes brings to light faults not otherwise discovered.

The question of how much money and time can be spent for the design of marginal checking and diagnostic circuitry was brought up. Replies emphasized that it is difficult to prejudge this in the design stage and that if designed-in the costs are difficult to separate later. The advantages of the use of error detecting and error correcting codes in the design of a computer were pointed out together with the fact that the cost of these is intimately interwoven into the over-all design.

A question inviting discussion of marginal testing methods for analogue computers drew no response.

Group Discussion on Crystal Diodes

RALPH J. SLUTZ†, *Moderator*

Introduction

A GROUP OF about seventy-five people met to discuss the reliability of crystal diodes. Dr. Slutz acted as discussion leader. The discussions are summarized in the five categories treated in the following paragraphs:

Life Testing

To launch the discussion, the moderator raised the question of the nature of the life characteristic of crystal diodes—whether the number still usable drops off suddenly at some particular life, whether the number falls off steadily with time as an exponential curve, or whether an asymptote is approached, with the remaining diodes appearing to have indefinitely long life. No one quoted carefully-controlled experiments to distinguish among these possibilities, but some opinions favored the last—a relatively long life after an initial period of "weeding out" faulty units. The question was then asked whether tests to a given percentage of the initial back current of a diode would indicate more rapid aging for high-resistance diodes than for low ones. Measurements were quoted which indicate that initially-high-back resistance units vary faster than those with lower initial back resistance.

There was some discussion of the methods of testing diodes for end of life. Since the resistance characteristic is so nonlinear, one speaker made a plea for measuring and talking about voltages and currents rather than resistances. It also was suggested that ac observations should be made, with oscilloscopic observation, since this gives a rapid check of the shape of the characteristic and shows changes of this shape during the observation period. The need was also mentioned for observing intermittent shorts or opens, particularly under vibration or tapping.

Transients in Diodes

Everyone seemed to have had trouble with either forward or backward transients. Storage and diffusion of charge carriers cause widely different volt-ampere characteristics from the normal dc characteristics, occurring in the microsecond range. One speaker said he had observed transient responses to be much improved by operating diodes at reduced currents and voltages.

† National Bureau of Standards, Washington, D. C.

Humidity

A good part of the discussion was about humidity effects on diodes. It was reported that diodes in a vacuum are very stable, but for practical units everyone agreed that a combination of high temperature and high humidity is likely to cause difficulty. The humidity effects show up more readily on cyclic tests than on continuous exposure to constant temperature and humidity. It was agreed that very careful sealing of the units can do much to reduce the effects of humidity, but it was stated that glass seals are not cure-alls—a ceramic can be as good as glass if special care is given to the impregnant, but active chemicals sealed within a glass unit can cause havoc.

Other Diodes

At one point the discussion was broadened to include other diodes. Some of these were gold-bonded germanium, and junction diodes of selenium, germanium, and silicon. Silicon diodes have in general a higher forward resistance, but a much higher back resistance than germanium. It was pointed out that selenium diodes deteriorate when left without applied back voltage. Experience was quoted in which selenium diodes carrying continuous forward current for several days suffered changes in the back characteristic of up to 1,000:1. Other experience was quoted in which selenium diodes were successfully operated for hundreds of hours at half the back voltage to which they had been originally formed. On the other hand, preliminary experience with silicon junction diodes indicates not only do they have very good high-temperature performance, but no deterioration was observed in several thousand hours of life.

Operating Experience

So far it has been difficult to interpret operating experience at a particular installation in quantitative terms of diode life. At one computer, only 230 out of 17,000 diodes have been replaced in 18 months of operation, but these were replaced as they were found to cause malfunction, and not as a result of repeated testing of all of the diodes in the computer. Oftentimes there are particular computer locations where diodes may deteriorate by orders of magnitude before causing trouble.

In general the discussion pointed out a need for more closely-controlled experience before we can say very much about the quantitative aspects of diode life.

Group Discussion on Technical Applications

G. W. PETRIE†, *Moderator*

THE GROUP discussion involved about 60 individuals representing both manufacturers and users of computing equipment.

Questions discussed were the following:

1. Are there better ways of coding machines for computing problems?
2. Where can one obtain mathematical service for analysis of machine operations such as determining the numerical code for a milling machine?
3. Is it more efficient for everyone to do his own coding for technical applications, or should a small group do the coding for all customers of the computing facilities?
4. What provisions must a programmer make for machine failure, power interruption, or other break in calculation?
5. In floating-decimal operations, what extreme ranges of exponents have been encountered?
6. How can one compare a digital computer with a differential analyzer for real-time control problems? Is there experience bearing on this subject?
7. What can be done to prevent customers from requesting excessive amounts of machine time due to unnecessarily large ranges of parameters?

The discussion revealed a wide range of opinion on almost all of the questions considered. Advantages of decentralized coding appear to be: The individual concerned retains complete control of his problem; he becomes acquainted with full-machine capabilities; he can accommodate the program to changes in the formulation of the problem; he avoids the communication delay of making his problem familiar to a coder. Advantages for a centralized coding group appear to be: More efficient coding because of an experienced coder familiar with all routines already available; the use of computing facilities which would otherwise be ignored because of the disinclination of engineers and researchers to do their own coding.

Extreme ranges for exponents of ten in floating-decimal operations cited to the group were -106 to $+8$ in an actual problem, while in another problem provision was made to include -236 to $+236$.

The discussion pointed directly to the fact that individual computing services adapt themselves readily to the needs of their sponsors instead of searching for a theoretically optimum mode of operation.

† International Business Machines, Washington, D. C.

Group Discussion on Magnetic-Tape Standards

ALLEN SHOUP†, *Moderator*

ABOUT THIRTY-FIVE persons attended an informal discussion regarding problems and points of interest pertaining to magnetic tape for use in the computer field. Several engineers explained some of their tape problems and the assembly in general very co-operatively offered helpful suggestions. There was considerable discussion regarding standardization of tape. It is impossible at this early date to set any standards, but various issues were submitted for consideration and these are being presented in this report. Everyone interested is invited to write the chairman regarding any of these subjects or any new items which they would like to discuss. After this information has been exchanged and collated, tentative standards will be formulated and distributed for comment. It is understood that although it will be some time before all the machines in the industry can become standardized, the present discussion may be of help to engineers starting new designs because they will be cognizant of the trend at this time.

Definitions

As a prelude to standardization, accurate definitions must be established. Definitions in the magnetic-tape field must reflect the thinking of a major portion of the engineers and manufacturers involved. The chairman would appreciate the opportunity to pursue this matter further by receiving comments and suggestions, both for items to be defined and associated definitions. Upon receipt and consolidation of this material, a summary report will be distributed.

Method of Test

Recommendations for standardization of tape factors must include the method by which factors are measured

† Shoup Engineering Co., Chicago, Ill.

experimentally. A standard cannot be established without specifying method of test, thus allowing industry to conduct its measurements with recognized procedure.

Recommendations for Standards

The establishment of standards of magnetic-tape parameters will be of mutual value to the industry. The standards should be in terms of variables which can be expressed in a quantitative manner. Among the factors which require standardization are the following:

Signal-to-Noise-Ratio

In addition to the properties of the tape, this ratio is a function of the frequency range, tape speed, tape weave, track width, type of head, and the way the mechanism holds the tape against the head. The tape should be measured under closely controlled conditions so that the results will reflect the properties of the tape rather than the properties of the equipment. The signal level must be specified in terms of the magnetic-saturation properties of the tape; it may be the signal which causes a specified percentage of harmonic distortion under standard conditions of bias frequency and tape speed. Standard methods of measuring noise level are available; however, care must be taken to exclude signal variations due to factors such as contact variations and flutter. Ideally, the signal-to-noise ratio for standard tapes should be of the order of 60 to 70 db.

Width of Tape

A great variety of tape widths are now being used. The engineers in attendance indicated the use of tapes with the following widths—$\frac{1}{2}$ inch, $\frac{5}{8}$ inch, $\frac{7}{10}$ inch, 1 inch, $1\frac{3}{4}$ inches, 2 inches, $2\frac{1}{2}$ inches, 3 inches, 4 inches, 35 millimeter, $17\frac{1}{2}$ millimeter, 16 mm single sprocket, 16 mm double sprocket. It would be advantageous to establish standard widths, such as $\frac{1}{2}$ inch, 1 inch, 2 inches, and 3 inches, so that the manufacturers could stock these widths. A reduction of the number of widths used on magnetic-tape recorders implies a curtailment of the number of track widths in use. There is evidence to support the concept that a change of track width can be accompanied by a change in useful packing factor to maintain constant the area of tape used in a particular application. In some instances, the tracks can be made wider, thereby allowing tape speed to be reduced to a point at which there could even be a saving in tape.

Track Widths

The group indicated that most of the tracks used are from 30 to 40 mils wide, but there are machines with track widths of 18, 25, 36, 50, 70, 88, 102 and 128 mils. During the discussion on this subject it appeared that a considerable number of the group were having trouble on the tape with very small bumps of emulsion (nodules) and with tape weave, with the result that pulses were missed. The chairman told the group that his company had overcome this difficulty completely, due mostly to the use of wider tracks (93 mils). Since 93 mils is the commercial standard this would be good standard to adopt.

Maximum Useful Tape Packing Factor

In general, the maximum useful tape-packing factor is limited by the design of the head. However, with continued research into increasing the packing factor a condition may arise making the tape a contributing factor in the packing-factor limitation. In the field of recording at video frequencies, it has been found that the type of magnetic material used can affect materially the response of the tape in the megacycle region. A standard test at audio and video frequencies will insure that the tape will have certain minimum capabilities at high frequencies.

Tape Factors Causing Flutter

Physical properties of the tape such as thickness, width, and stiffness variations, will in general cause flutter and tape weave. Expert mechanical design of the transport mechanisms can reduce the effect of these variations, but the establishment of standards on the physical properties of the tape which contribute to flutter could benefit manufacturers of all classes of tape recorders.

Mechanical Characteristics

The group expressed an interest in mechanical features of the tape. For some of the parameters, suggestions were received from the floor, as noted in the list below.

(a) Tensile strength.
(b) Number of passes through machine.
(c) Tolerance on width of tape.
(d) Tolerance on the edges of the tape conforming to a straight line rather than a slight curvature.
(e) Length of reel—2,400 feet was suggested.
(f) Thickness of tape and tolerance on the thickness.
(g) Properties of the surfaces of the tape.
(h) Temperature characteristics (flexibility at low temperatures; tensile strength, surface conditions, etc. at high temperatures).
(i) Stiffness of tape.
(j) Emulsion characteristics (to what degree the emulsion tends to scratch or pile up at some spot in the mechanism).
(k) Humidity characteristics of tape.

Tape Storage

Some discussion centered about the length of time that tape records would be required to be kept. The concensus was that a tape need retain its magnetization and mechanical characteristics for at least one year but not more than two.

The specifications of the National Bureau of Standards Computer tape system as described by Dr. S. N. Alexander, as well as those of several other groups represented at the meeting, were kindly offered to the committee by the representatives present. The chairman would welcome additional specifications from groups wishing to encourage the standardization program.

Group Discussion on Commercial and Industrial Applications of Computers

W. F. FRESE†, *Moderator*

THE DISCUSSION at this session, attended by about 100 people, related almost entirely to the application of electronic techniques to government and commercial accounting systems. The chairman first summarized the implications of the government's joint accounting improvement program. In the past five years, the emphasis has shifted from uniformity of practice in the various agencies to greater decentralization, with details handled and responsibility located at or near the point of origin of information. The amount and type of mechanization has been put up to individual agencies; central authorities co-operate with all agencies, and are constantly looking for methods that reduce duplication of effort. The concept of what accounting should accomplish has broadened, and there is now less interest in historical record-keeping, and much more on provision and integration of the phases required (financial, cost, inventory, etc.) to provide information needed for current action of management.

There was considerable discussion of the procedures and operations involved in the payment and reconciliation of Government punched-card checks (of which nearly 300,000,000 are issued per year) and the intensive studies being made to consolidate and integrate payment and reconciliation operations through application of electronic computers and related techniques.

This example elicited questions as to difficulties with multilation of punched-card checks. Initially multilation ran as high as 10 per cent, but continuing educational programs keep it below 2 per cent, which is entirely acceptable. Similar experience was cited for the punched-card money orders now used by the Post Office. Reference was also made to the work of the First National Bank of Chicago in this area; card presses developed for them, which reshape creased or folded cards, have pushed below 1 per cent the fraction of cards unacceptable to machines. Development of magnetic inks, as a substitute for punching, was also mentioned.

The discussion then turned to difficulties in obtaining management acceptance of full electronic accounting systems. It was emphasized that their application requires breaking through present organization lines to achieve an economical integrated system by consolidation of related operations. Some speakers felt the necessary functional reorganization would be hard to sell management, but demonstrable economy would ultimately win out; others felt top management easily sold, though admitting it would be hard at lower levels.

The next question related to machine accuracy actually necessary for accounting applications. Accountants present indicated willingness to evaluate this by realistically comparing over-all costs of the calculated risk approach with that of complete accuracy, wherever the comparison can be done with reasonable realibility.

Inquiry was made as to the acceptability of magnetic tape as a primary accounting record, and its status as legal evidence. The discussion indicated that the specific question could not yet be answered; no instance of use of tape as primary record was brought to light. However, many situations were cited, where tape could serve as a convenient and economical means of locating a needed primary record or information derived from it, with the record itself preserved in another form of unquestioned legal status. Several of these involved frequent searching of files of millions of items. The use by the Army, of decks of punched cards as substitutes for written reports on the status of supply inventories, was also noted; in this connection, the desirability of radio and land-line communication links designed to tie together the parts of geographically distributed data-processing systems was brought out.

A number of questions related to the availability of smaller, less expensive electronic units suitable for small business, and of information on them and on their proper utilization. These brought forth a number of answers. One was that many companies are building or developing computers in the $50,000 to $100,000 class. A second was that it does not take as much to justify the cost of a large computer for accounting operations as is often thought; specifically, that a large computer operating as much as ½ shift per day will frequently prove economical. Third was that a computer of any size is essentially a time-sharing device, and thus admirably adapted to serving a number of small companies who jointly support it. In this respect, several small users would be in much the same position as a number of departments of a large organization; the additional advantage that nondisclosure of data to subsequent users is much easier to ensure than with (say) punched-card equipment was also pointed out. The listing of companies and their developments in "Computers and Automation" and the ONR Digital News Letter, henceforth to be published in the ACM magazine, and the existence of consulting engineering firms equipped to serve prospective accounting users, was also mentioned.

The need for more effective communication between accountants and computer engineers was mentioned repeatedly during the session; while there was some difference of opinion as to means of improving the mutual understanding between the users and the manufacturers, there was general agreement that professional societies could make a great contribution to the effective application of electronic computer techniques in business by clearing the channels of communication.

† U. S. General Accounting Office, Washington, D. C.

After-Luncheon Remarks

H. R. J. GROSCH†

I CANNOT help feeling flattered by our chairman's introduction. In referring to me as an expert, however, he put me under the obligation of reminding you of the definition: "X" stands for an unknown quantity, and "spurt" is a drip under pressure!

Our field has had a most remarkable growth in the last half decade. Counting exhibitors, paid registrants, and guests, we have about fourteen hundred people at this three-day meeting. Those of us present at this luncheon would have been considered quite a sizable computer group five or six years ago. It is probably safe to say that there is no other field of technology which has grown so rapidly as our own. And I include even television, at least technically, although we have to admit that so far they have tapped a slightly larger mass market!

One of the things that is most encouraging in the computer field is the strong increase in interest we can now observe in the detailed control of manufacturing processes and in factory automation. The first involves both operations research and the mechanization of clerical operations; the latter brings us close to the real-time control problem, since it envisages the actual control of tools by digital programming devices.

Earlier in this meeting we heard a paper by Vernon Weihe on another aspect of real-time control, namely, the mechanization of the air-traffic-control problem with a digital computer as the central logical organ. There is little doubt that this class of problems, requiring accurate decisions over a complex field of information at more than human speeds, represents tremendous application potential.

There is still another field in which we may expect major developments: a short label might be "information retrieval." The insurance companies, from whom we have heard at this meeting, the magazine subscription offices, patent and legal searchers, technical libraries, and many other agencies are faced daily with searching problems which tempt us by their size, complexity, and even their commercial importance.

We are so prosperous now, and of such interest to the business and popular press, that it has become quite difficult even for applications men to keep track of all these new areas of interest. A new class of experts—or perhaps in view of our chairman's introduction, I had better use the term "synthesists"—who are interested in the broad application of high-speed computers in all fields, has recently been given considerable prominence. I have named this species in honor of its earliest and best known proponent, but in spite of this, I am sure that the species is valuable and will continue to flourish.

I should now like to pass to a more serious topic. We have had increasing support, at this and other recent meetings of the hardware and application fraternities, for the views that I used to express back in 1948 and 1949 on the growing shortage of personnel. There is no question about successful development and use of high-speed information-processing equipment depending almost entirely on the availability of high-class systems engineers, experienced numerical analysists, and imaginative programmers. Educating these people, and giving them working experience, has been and continues to be one of the most difficult bottlenecks in the expansion of our field. Some universities have made a beginning, and even their mathematics departments no longer shy away from our area, but unfortunately outstanding institutions like MIT have a bad habit of consuming most of their own output! The problem is now taken seriously by everyone, however, and I am sure that both university and manufacturer training programs will grow in size and effectiveness.

It is now time for us to think seriously about another area which requires support for its expansion. This support must come from industrial users and from computer and computer component manufacturers; the current emphasis on economy in government, which unfortunately appears to hit the most worthwhile research projects first, makes it unlikely that federal support in this field will be continued; certainly such support will not be increased. This field may be conveniently labeled as research programming. It includes such efforts as Grace Hopper's compiler routines, Tony Oettinger's conditioned-reflex-learning study, and Strachey's checker playing. It also should include our early adventures in minimum latency drum-computer programming, programmed redundancy checks, numerical methods such as bivariate function approximation, and paper studies of the economics of simulation. Other topics at a similar research level will occur to all of you, even those most concerned with componentry.

Research programming requires support for lengthy periods of time. Commercially, valuable results are not to be guaranteed, although they almost certainly will appear. The analyst who is to follow an idea which may later generate an entirely new field of application will frequently not be interested in immediate commercial results. I think of Shannon's maze-solving machines as an outstanding example. Yet, like basic research in older fields like solid-state physics or biochemistry, it is tremendously important that we have at least a small

† Aircraft Gas Turbine Div., Gen. Elec. Co., Cincinnati, Ohio.

segment of our technical people working at long range problems of their own choosing. For this class of investigation I should like to appeal strongly for at least two sources: small research groups within the organizations of the large manufacturers in our field, and industrial and federal—perhaps National Science Foundation—support of university efforts. All of you recognize prototypes of both classes, but I am calling for at least a tenfold expansion in the next three or four years, if we are to utilize the full potentialities of the new generations of computing equipment as they become available. Research in this programming area can be tremendously stimulating to our entire field of interest, and over a long period of time can pay big dividends to our increasingly technical civilization, but support is required now, on an increasing scale, and with a wider geographical distribution.

I wish that every one of you who has connections with an imaginative management would consider soliciting its support for this sort of work, either within your own company or branch of government or by support of specialized university fellowship and research programs. In this way we can all help keep the tremendous expansion of computer applications at its current height.