MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A
SOFTWARE TEST HANDBOOK

Boeing Aerospace Company

Edward Presson
This report has been reviewed by the RADC Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-84-53, Vol I (of two) has been reviewed and is approved for publication.

APPROVED:  
FRANK S. LaMONICA  
Project Engineer

APPROVED:  
RAYMOND P. URTZ, JR.  
Acting Technical Director  
Command & Control Division

FOR THE COMMANDER:  
JOHN A. RITZ  
Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC (COEE) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document requires that it be returned.
**REPORT DOCUMENTATION PAGE**

1. **REPORT SECURITY CLASSIFICATION**
   - UNCLASSIFIED

2a. **SECURITY CLASSIFICATION AUTHORITY**
   - N/A

2b. **DECLASSIFICATION/DOWNGRADING SCHEDULE**
   - N/A

3. **DISTRIBUTION/AVAILABILITY OF REPORT**
   - Approved for public release; distribution unlimited

4. **PERFORMING ORGANIZATION REPORT NUMBER(S)**
   - N/A

5. **MONITORING ORGANIZATION REPORT NUMBER(S)**
   - RADC-TR-84-53, Vol I (of two)

6a. **NAME OF PERFORMING ORGANIZATION**
   - Boeing Aerospace Company

6b. **OFFICE SYMBOL**
   - Engineering Technology

7a. **NAME OF MONITORING ORGANIZATION**
   - Rome Air Development Center (COEE)

7b. **ADDRESS (City, State and ZIP Code)**
   - Seattle WA 98124

8a. **NAME OF FUNDING/SPONSORING ORGANIZATION**
   - Rome Air Development Center

8b. **OFFICE SYMBOL**
   - COEE

8c. **ADDRESS (City, State and ZIP Code)**
   - Griffiss AFB NY 13441

9a. **NAME OF RESPONSIBLE INDIVIDUAL**
   - Edward Presson

9b. **TELEPHONE NUMBER**
   - (315) 330-3977

10. **SOURCE OF FUNDING NO.**

11. **TITLE (Include Security Classification)**
   - SOFTWARE TEST HANDBOOK

12. **PERSONAL AUTHOR(S)**
   - Edward Presson

13a. **TYPE OF REPORT**
   - Final

13b. **TIME COVERED**
   - From Mar 62 to Sep 83

14. **DATE OF REPORT (Yr., Mo., Day)**
   - March 1984

15. **PAGE COUNT**
   - 60

16. **SUPPLEMENTARY NOTATION**
   - N/A

17. **COSSATI CODES**

<table>
<thead>
<tr>
<th>FIELD</th>
<th>GROUP</th>
<th>SUB.GRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>02</td>
<td>Testing</td>
</tr>
<tr>
<td>39</td>
<td>02</td>
<td>Software</td>
</tr>
<tr>
<td>39</td>
<td>02</td>
<td>Computer Programs</td>
</tr>
</tbody>
</table>

18. **SUBJECT TERMS (Continue on reverse if necessary and identify by block number)**

19. **ABSTRACT (Continue on reverse if necessary and identify by block number)**

The purpose of the Software Test Handbook effort was to provide Air Force software developers with guidelines and methodology for the effective use of higher order language (HOL) software testing techniques and in the selection of automated tools for the testing of computer programs.

The effort resulted in a two volume final technical report. This report, Volume I - Final Technical Report, describes the total contractual effort including a project overview, summary for each of three technical tasks, and a bibliography. Volume II, the Software Test Guidebook, contains the guidelines and methodology resulting from the effort.

20. **DISTRIBUTION/AVAILABILITY OF ABSTRACT**
   - UNCLASSIFIED

21. **ABSTRACT SECURITY CLASSIFICATION**
   - UNCLASSIFIED

UNCLASSIFIED

**DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.**

UNCLASSIFIED
PREFACE

This document is the final technical report, CDRL item A005, that describes the results of the three tasks involved in developing the Software Test Guidebook. This report was prepared in accordance with the statement of work, contract F30602-82-C-0059. It summarizes the activities performed for the Rome Air Development Center by Boeing Aerospace Company in Kent, Washington.

In task 1 of this contract, a survey was made to determine current software testing practices used in the five major Air Force missions. The survey was supplemented with five site visits. Task 2 evaluated current state-of-the-art software testing techniques and test tools and incorporated this information into tables. Task 3 designed and prepared the handbook.

This report is in two volumes: the technical report that describes the total contractual effort and the Software Test Guidebook that describes the methodology for selecting testing techniques and test tools. The technical report comprises a project overview in section 1.0; summaries of tasks 1, 2, and 3 in sections 2.0, 3.0, and 4.0; and a bibliography in section 5.0. The Software Test Guidebook is divided into six sections. It is designed to assist Air Force software developers in using higher order language software testing techniques and in selecting automated tools to test computer programs.
TABLE OF CONTENTS

1.0 PROJECT OVERVIEW 1
   1.1 Background 1
   1.2 Project Summary 2
   1.3 Scope of Effort 2
   1.4 Brief Description of the Guidebook 3
   1.5 Outline of Investigation 6
      1.5.1 Task 1, Programming Environment Survey 6
      1.5.2 Task 2, Evaluation of Testing Techniques 7
      1.5.3 Task 3, Preparation of Software Test Guidebook 8

2.0 SUMMARY OF TASK 1 9
   2.1 Chronology of Activities 9
   2.2 Survey Methodology 11
   2.3 Survey Findings 13
      2.3.1 Observations 13
      2.3.2 Air Force Site Summaries 16

3.0 SUMMARY OF TASK 2 19
   3.1 Evaluation of Testing Techniques 19
   3.2 Task 2 Considerations 20
   3.3 Selection of Tool Taxonomy 23
   3.4 Software Environments Characteristic of USAF Missions 26

4.0 SUMMARY OF TASK 3 28
   4.1 Guidebook Development 28
   4.2 Preparation of Text 28
   4.3 Preparation of Graphic Materials 29
   4.4 Editing and Review 29

5.0 BIBLIOGRAPHY 30
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Guidebook Organization</td>
<td>4</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>Armament Division</td>
<td></td>
</tr>
<tr>
<td>ADC</td>
<td>Aerospace Defense Center</td>
<td></td>
</tr>
<tr>
<td>ADP</td>
<td>automatic data processing</td>
<td></td>
</tr>
<tr>
<td>AFTI</td>
<td>advanced fighter technology integration</td>
<td></td>
</tr>
<tr>
<td>ALCM</td>
<td>air-launched cruise missile</td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>Aeronautical Systems Division</td>
<td></td>
</tr>
<tr>
<td>ATE</td>
<td>automatic test equipment</td>
<td></td>
</tr>
<tr>
<td>ATO</td>
<td>Air Tasking Order</td>
<td></td>
</tr>
<tr>
<td>BITS</td>
<td>built-in test</td>
<td></td>
</tr>
<tr>
<td>CAFMS</td>
<td>Computer Assisted Air Force Management System</td>
<td></td>
</tr>
<tr>
<td>CCPDS</td>
<td>Command Center Processing and Display System</td>
<td></td>
</tr>
<tr>
<td>CINCSAC</td>
<td>Commander-in-Chief Strategic Air Command</td>
<td></td>
</tr>
<tr>
<td>CITS</td>
<td>central integrated test systems</td>
<td></td>
</tr>
<tr>
<td>COM</td>
<td>computer output microfiche</td>
<td></td>
</tr>
<tr>
<td>CPCI</td>
<td>computer program configuration item</td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>central processor unit</td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>command, control, and communications</td>
<td></td>
</tr>
<tr>
<td>CSOC</td>
<td>Consolidated Space Operations Center</td>
<td></td>
</tr>
<tr>
<td>DGZ</td>
<td>designated ground zero</td>
<td></td>
</tr>
<tr>
<td>DT&amp;E</td>
<td>development, test and evaluation</td>
<td></td>
</tr>
<tr>
<td>ELINT</td>
<td>electronic intelligence</td>
<td></td>
</tr>
<tr>
<td>HDM</td>
<td>hierarchical design methodology</td>
<td></td>
</tr>
<tr>
<td>HOL</td>
<td>higher order language</td>
<td></td>
</tr>
<tr>
<td>ICBM</td>
<td>intercontinental ballistic missile</td>
<td></td>
</tr>
<tr>
<td>IDHS</td>
<td>intelligence data handling system</td>
<td></td>
</tr>
<tr>
<td>IUS</td>
<td>inertial upper stage</td>
<td></td>
</tr>
<tr>
<td>IV&amp;V</td>
<td>independent verification and validation</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>JINTACCS</td>
<td>Joint Interoperable Tactical Air Command and Control System</td>
<td></td>
</tr>
<tr>
<td>JSCS</td>
<td>Joint Strategic Connectivity Staff</td>
<td></td>
</tr>
<tr>
<td>JSTPS</td>
<td>Joint Strategic Target Planning Staff</td>
<td></td>
</tr>
<tr>
<td>LRU</td>
<td>line replaceable unit</td>
<td></td>
</tr>
<tr>
<td>MATE</td>
<td>Modular Automatic Test Equipment (program)</td>
<td></td>
</tr>
<tr>
<td>NMCC</td>
<td>National Military Command Center</td>
<td></td>
</tr>
<tr>
<td>O&amp;S</td>
<td>operations and support</td>
<td></td>
</tr>
<tr>
<td>OFP</td>
<td>operational flight programs</td>
<td></td>
</tr>
<tr>
<td>OPR</td>
<td>Office of Primary Responsibility</td>
<td></td>
</tr>
<tr>
<td>OT&amp;E</td>
<td>operational test and evaluation</td>
<td></td>
</tr>
<tr>
<td>PROM</td>
<td>programmable read-only memory</td>
<td></td>
</tr>
<tr>
<td>SAC</td>
<td>Strategic Air Command</td>
<td></td>
</tr>
<tr>
<td>SCF</td>
<td>Satellite Control Facility</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>Space Division</td>
<td></td>
</tr>
<tr>
<td>SDL</td>
<td>software development laboratory</td>
<td></td>
</tr>
<tr>
<td>SILTF</td>
<td>System Integration Laboratory and Test Facility</td>
<td></td>
</tr>
<tr>
<td>SIOP</td>
<td>Single Integrated Operational Plan</td>
<td></td>
</tr>
<tr>
<td>SLBM</td>
<td>submarine-launched ballistic missile</td>
<td></td>
</tr>
<tr>
<td>SOLARS</td>
<td>SAC On-Line Analysis and Retrieval System</td>
<td></td>
</tr>
<tr>
<td>SPO</td>
<td>System Program Office</td>
<td></td>
</tr>
<tr>
<td>SREM/REVS</td>
<td>Software Requirements Engineering Methodology/Requirement Engineering Validation System</td>
<td></td>
</tr>
<tr>
<td>SRU</td>
<td>shop replaceable unit</td>
<td></td>
</tr>
<tr>
<td>TAC</td>
<td>Tactical Air Command</td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Tactical Air Control Center</td>
<td></td>
</tr>
<tr>
<td>TACS</td>
<td>Tactical Air Control System</td>
<td></td>
</tr>
<tr>
<td>TRD</td>
<td>Test Requirements Document</td>
<td></td>
</tr>
<tr>
<td>TRICOMS</td>
<td>Triad Computer System</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>TTY</td>
<td>teletypewriter</td>
<td></td>
</tr>
<tr>
<td>USAF</td>
<td>United States Air Force</td>
<td></td>
</tr>
<tr>
<td>UUT</td>
<td>unit under test</td>
<td></td>
</tr>
<tr>
<td>V&amp;V</td>
<td>verification and validation</td>
<td></td>
</tr>
<tr>
<td>WWMCCS</td>
<td>Worldwide Military Command and Control System</td>
<td></td>
</tr>
</tbody>
</table>
1.0 PROJECT OVERVIEW

1.1 BACKGROUND

The testing and operational support of computer programs continue to be critical information processing problems facing the Air Force. Substantial resources in terms of funds and personnel are continually applied during the software development life cycle for testing and maintaining computer programs. However, in many instances the application of these resources does not achieve the desired result; that is, programs thought to be correct may suddenly produce erroneous outputs during operational use.

Traditionally, testing, verification, and validation of software have been a largely manual process. For example, test drivers and data are usually prepared manually and the test results manually interpreted. A wide variety of additional software testing and verification techniques has been developed in recent years, many of which can and have been implemented in automated software tools. While the effectiveness of manually applied techniques can vary considerably with the skill with which they are applied, many software tools can be highly consistent and reliable in their ability to detect the errors for which they were designed. These software tools can potentially improve the quality of testing and at the same time reduce the manual effort required to test a software system.

Typical testing activities that occur in the software development life cycle include unit testing, component testing, integration testing, system-level and acceptance testing, and maintenance testing (retesting). These activities, which may differ slightly from one development environment to another, are each associated with different objectives and impose specialized requirements on the testing task. As a result, certain testing techniques may be more applicable to one activity than to another.

In many cases, while the techniques implemented by the tools have proved effective, the use of these tools in military software development and support environments has been lacking. In part, the low utilization is because managers and engineers are not well informed about the availability of tools and techniques, their usefulness and effectiveness, and how they can be integrated properly into specific development and support environments. While a number of regulations and guidebooks for software development have been
prepared in the past, most have dealt with providing an understanding of the software development life cycle, and little emphasis has been placed on using software testing tools and techniques in the life cycle.

1.2 PROJECT SUMMARY

The purpose of the Software Test Guidebook project is to provide Air Force software developers with a guidebook to guide them in the effective use of higher order language (HOL) software testing techniques and in the selection of automated tools for the testing of computer programs. The guidebook specifies guidelines and methodologies for understanding and applying automated state-of-the-art testing techniques in various types of Air Force software development and support environments. This effort will foster the transfer of advanced software testing technology to the Air Force user community. Air Force software developers may find the results of this effort useful in supplementing existing regulations and guidebooks.

1.3 SCOPE OF EFFORT

Guidelines and methodologies were developed that describe the proper use of advanced software testing technology during the development of computer application software for the five primary Air Force missions (armament, avionics, command, control, and communications (C^3), missile/space, and mission/force management). The guidelines and methodologies pertain to those testing activities of the software development life cycle that follow the beginning of actual program coding. Representative Air Force software sites were visited and analyzed to determine typical characteristics of application software and environments in which application software is developed and maintained. Characteristics of advanced state-of-the-art software testing technology were extracted from the literature and from available software tool surveys. The guidelines and methodologies developed have been provided in the form of a guidebook.

Guidelines and methodologies were developed for the selection of state-of-the-art software testing techniques in the computer program development life cycle; that is, development test and evaluation (DT&E), operational test and evaluation (OT&E), and verification and validation (V&V), as defined in AFR 80-14 and AFR 800-14, with the following constraints:
a. This effort covered only the coding and checkout phase, test and integration phase, and operation and support (O&S) phase. Further, for the O&S phase, the investigation was limited to the support aspects (i.e., coding, checkout, and retesting of modifications to deployed computer programs).

b. This effort considered not only those testing techniques pertinent to the development of operational computer programs for the five primary Air Force missions, but also the use of those testing techniques in the development of auxiliary software programs (e.g., simulators and data analysis programs) used to test and support the operational computer programs.

1.4 Brief Description of the Guidebook

The principal purpose of the guidebook is to select state-of-the-art testing techniques for organic Air Force software testing. In addition, the guidebook can be used for preparation of a Statement of Work and for evaluation of proposals. All three applications use the same table-driven methodology to determine appropriate software testing techniques.

The guidebook is designed to permit a user to determine appropriate software test techniques, starting with the knowledge of which USAF mission is being supported. Using the unique guidebook appendix to his mission, the user can identify the specific software types to be tested. With this knowledge, the user will be led to the generic software type corresponding to the software function at hand. With the generic software type (from the guidebook appendix) and testing confidence level, the user can go to another table to find the appropriate software testing techniques for his unique situation.

The table-driven methodology provides three paths for technique selection. The first path is based on the kind of software being tested and the testing confidence level required. The second path is based on the test phases and test objectives, and the third path is based on detection of specific software error types.

The contract was titled "Software Test Handbook," but "Handbook" was changed to "Guidebook" to better reflect the purpose, which was to assist in the selection of testing techniques, not to rigidly define them.
Figure 1. Guidebook Organization
The Software Test Guidebook comprises six major sections and five appendices, as shown in figure 1. A brief summary of its contents is found in the following paragraphs.

Section 1.0 states the objectives of the guidebook, describes its outline and content, and discusses its applications.

Section 2.0 presents a compact set of instructions, guidelines, and tables for selecting software testing techniques. It also includes sections on the selection of software support tools and on test completion criteria.

Section 3.0 contains a list of the major catalogs that provide information on automated software tools. It has tables that provide a cross reference to three tool catalogs for determining the availability of existing software tools that support the techniques selected by the guidebook user.

Section 4.0 defines the terms used in the taxonomy of testing techniques and gives a detailed description of state-of-the-art testing techniques. These descriptions discuss the technique and related considerations such as cost, user training, and hardware requirements.

Section 5.0 discusses the software life cycle, software acquisition cycle, and the normal phases of testing, as defined in AFR 80-14 and 800-14. This section is a supplement to the guidebook and does not replace the Air Force regulations.

Section 6.0 has several model statement of work (SOW) paragraphs that may be used as prototypes by Air Force acquisition managers in preparing a request for proposal (RFP).

The appendices describe five Air Force mission areas: armament; avionics; command, control, and communication (C³); missile/space; and mission/force management. Each appendix lists software functions characteristic of the computer programs developed within that mission. These functions are each assigned a software category number.
1.5 OUTLINE OF INVESTIGATION

The Software Test Guidebook development comprised three major tasks, which are outlined in sections 1.5.1 through 1.5.3.

1.5.1 Task 1, Programming Environment Survey

An investigation to analyze programming characteristics and the development and support environments of typical application software for the five primary Air Force missions was conducted in task 1.

Within each primary mission, differences in the application software that impact the most appropriate testing and verification strategy were investigated. The following characteristics were addressed:

- Operating instructions and regulations used to standardize programming and testing methods and to document the testing process.
- Level of robustness (fault tolerance).
- Timing and synchronization requirements; real-time processing constraints.
- Distributed and centralized processing configurations.
- Application program processing requirements.
- Type and level of HOL programming languages used.
- Level of support and testing provided by compilers.
- Software testing tools and techniques and debugging aids typically used.
- Code complexity (data structures, control structures, data types).
- Real and non-real-time applications.
- Where coding and testing are performed (type of machine).
- Availability and abundance of memory, central processor unit (CPU), and disk storage.
- Availability of input-output (I/O) to an external storage media.
- Computer host and target relationship.
- Use and non-use of environment simulators on host computer.
- Batch and interactive mode of use.
- Size and number of modules in a typical application and sizing criticality.
- Evolutionary requirements of a typical software application.
- Type and criticality of software application (or its components).
- Mission-imposed time constraints for software modification.
- Testing requirements included in statements-of-work for contracted software development.

To collect characteristic data representative of the applications as specified in the preceding list, the following Air Force software development sites were visited:

- Tactical Air Command Headquarters, TAC/ADY, Langley AFB, Virginia.
- Aeronautical Systems Division, ASD/EN, Wright-Patterson AFB, Ohio.
- Space Command, SPACECOM/KRS, Peterson AFB, Colorado.
- Space Division, SD/AGM, Los Angeles AFS, California.
- Strategic Air Command Headquarters, SAC/AD, Offutt AFB, Nebraska.
- Armament Division, AD/SDEE, Eglin AFB, Florida.

A questionnaire was prepared and distributed to each of the designated sites prior to the scheduled visit. It addressed collecting current data pertinent to the previously listed application software and development environment characteristics. Results of this task were reported in the task 1 Interim Technical Report, CDRL item A003, dated February 1983, and are summarized in section 2.0. These results were also used to prepare appendices A through E of the Software Test Guidebook to describe the five primary Air Force mission areas. The P/M Group (1546 Marsetta Drive, Dayton, Ohio 45432) was responsible for the survey of AD and ASD.

1.5.2 Task 2, Evaluation of Testing Techniques

State-of-the-art HOL software testing techniques most appropriate to the individual testing and verification phases of each type of Air Force application software and its associated development and support environment were determined. A correlation of test requirements with those techniques most suited for verifying them was provided.

Examples of such requirements include module interface checking, timing and synchronization, and fault tolerance.
To determine the applicability of testing techniques to particular Air Force environments, the following characteristics were considered:

a. Performance analysis capability.
b. Error detection and location capability.
c. Side effects and benefits (e.g., computer program documentation).
d. Cost, schedule, and benefits impact.
e. Management impact.
f. Training required (user and maintenance training).
g. Usage constraints (machine, language, and operating system).
h. Typical computer resources required (e.g., memory and CPU time).
i. Level of human interaction required.
j. Usefulness of technique in supporting modern programming practices such as structured programming and structured walkthroughs.

Software tool taxonomies and surveys were used to define and identify software testing tools and techniques, and their typical performance characteristics. Results of this task are reported in section 3.0.

1.5.3 Task 3, Preparation of Software Test Guidebook

Guidelines and methodologies for using the applicable techniques during DT&E, OT&E, and V&V for each Air Force mission area were developed and compiled into a guidebook. The organization of the guidebook was designed around the table-driven methodology. The design goal was to produce a guidebook that would be easy to use for both the first-time reader as well as the experienced guidebook user. To achieve this goal, the guidebook was organized so that its use would appear intuitively right to the new reader, but arranged so the experienced user could accomplish his task efficiently. The guidebook organization and usage was illustrated with a series of figures to further clarify the structure. The text was written by software engineers and reviewed by a technical editor; this process went through several iterations. Additionally, the contract technical monitor provided many comments on early drafts as the guidebook evolved.
2.0 SUMMARY OF TASK 1

2.1 CHRONOLOGY OF ACTIVITIES

The following list of activities were completed in task 1.

a. The Software Test Handbook contract was awarded on March 3, 1982.

b. A kickoff meeting was held at RADC in Rome, New York, on March 25, 1982, where plans for the task I activities were discussed and a detailed proposal was presented for the questionnaire outline, scope, and coverage.

c. Air Force site representatives were contacted by telephone to familiarize them with the survey process; this activity took place during the week of April 17, 1982.

d. A pilot survey was conducted in May 1982, using the ASAT Missile Guidance Computer Program. This pilot survey was followed by interviews of the individuals who completed questionnaire forms in order to determine the time required to complete questionnaire sections and identify any deficiencies. Comments received from the pilot survey were incorporated into a revision to the questionnaire.

e. The draft questionnaire form was distributed to eight Air Force site representatives for their review and evaluation on May 14, 1982. Completion of this review required 1 month.

f. The questionnaires were completed by the following Air Force sites and returned by September 1982.
   1. Armament Division, AD/ENEC, Eglin AFB, Florida.
   3. Aeronautical Systems Division, ASD/EN, Wright-Patterson AFB, Ohio.
   4. Space Division, SD/AGM, Los Angeles AFS, California.
   5. Tactical Air Command, TAC/ADY, Langley AFB, Virginia.
Air Force onsite survey visits were conducted by Boeing and P/M Group personnel at six sites on the following dates:

1. Space Division (SD), Los Angeles. September 15 through 19, 1982, (P/M Group).


4. Tactical Air Command (TAC), Langley AFB. September 23 through 24, 1982, (Boeing).

5. Strategic Air Command (SAC), Offutt AFB. October 4 through 5, 1982, (Boeing).

6. Armament Division (AD), Eglin AFB. October 7 through 8, 1982, (Boeing).

At the four sites surveyed by Boeing, briefings were presented to Air Force personnel from all participating organizations. The presentations covered the purpose of the guidebook, the approach for using the guidebook, examples of some of its tables, and a taxonomy for software test techniques and tools developed for the guidebook.

Air Force site summaries were prepared during November 1982. These summaries were derived from a composite of the data obtained from the questionnaire responses and onsite visits.

An oral presentation covering a summary of task I activities took place at RADC on February 1, 1983. Representatives of all participating Air Force organizations were invited by RADC to attend this review.
2.2 SURVEY METHODOLOGY

The scope of the survey was well defined by the statement of work for the contract. The tasks to implement the survey required defining a general strategy for effectively obtaining the required information from participating sites. An underlying assumption was that the principal problems would be (1) limited availability of personnel to provide survey data and (2) determination of how to select the target software for survey coverage, so that the results would provide representative data; survey forms for recording data must be compatible with the level of the software. It would have been desirable to select classes of software for conducting the survey for each site. However, this was not practical because (1) the classes at the various sites were not known and (2) classification of software for site and mission areas was considered a necessary product derived from the survey, rather than an input to it.

It was concluded that the survey questionnaires should be directed toward individual computer program configuration items (CPCI). Each site representative was requested to select between four and eight computer programs, typical of their software, for the survey. This number was a compromise between a desire for a more statistically valid sampling and the significant amount of time needed to complete the questionnaires. The pilot survey participants reported that it took them approximately 3.5 hours to fill out one questionnaire.

Another consideration in formulating the questionnaire was whether the basic information should be obtained in person-to-person interviews or by a mailed-in form, followed up by a more generalized interview covering the total site software development and maintenance environment. It was concluded that the latter alternative would be a more efficient use of contract funds, permit greater flexibility in the nature of the onsite interviews, and help minimize bias.

Once the basic approach for the survey was defined—a stand alone questionnaire, followed by an unencumbered site visit—it was necessary to establish the format of the questionnaire to provide the needed information. Two general approaches were considered: one emphasized essay questions, the other emphasized multiple-choice questions. It was concluded that essay questions would be most desirable for person-to-person interviews, while multiple choice would be best for the mail-in-form approach. Therefore,
multiple-choice questions were selected for the questionnaire, despite the fact that they would be more difficult to prepare. Essay questions were considered too unstructured for the survey.

The questionnaire form was divided into six independent sections. The rationale for this was twofold: (1) to encourage qualified personnel to provide information in areas of their expertise and (2) to provide the capability to distribute the effort among several participants, rather than encumbering a single person. One large questionnaire was considered more likely to be deferred in the work scheduling than a sectionalized one. Many of the questionnaires were completed by more than one person, but rarely by more than three. The six sections of the questionnaire were as follows:

a. Part 1, General Software Information. Acquires information concerning the size and the technical and contractual requirements for a computer program. This part was intended to be completed by individuals such as program managers or administrative officers with project-level visibility.

b. Part 2, Development and Maintenance Environment. Acquires information about processing capabilities and development aids and tools used to develop, support, and maintain a computer program. This was intended to be completed by individuals such as project software managers with detailed knowledge of the software development and maintenance environment.

c. Part 3, Software Characteristics. Acquires information about the technical characteristics of a computer program, including design, code and data characteristics, and complexity. This part was intended to be completed by individuals such as lead programmers with detailed knowledge of the computer program.

d. Part 4, Software Tools. Acquires information about automated tools used for error detection and analysis and testing, including compiler-based tools and static and dynamic analyzers. This part was intended to be completed by individuals such as lead programmers or lead test engineers with general knowledge of the use of software tools. Ratings of tools used were intended to be done by tool users.

e. Part 5, Software Test Methods. Acquires information about testing activities for a computer program, including rationale for test cases and test completion criteria.
This part was intended to be completed by individuals such as software test engineers with in-depth knowledge of software test activities and error reports.

f. Part 6, Software Error Categories. Acquires information about types of errors detected during initial development or maintenance. This part was intended to be completed by individuals such as software engineers, test engineers, configuration management personnel, or quality assurance personnel who are responsible for recording or maintaining software error data.

2.3 SURVEY FINDINGS

This section represents summaries of the characteristics and nature of software development, maintenance, and testing practices of each of the sites. The findings are derived from a composite of data collected from the questionnaire forms and by onsite interviews. Generally, the interviews were conducted with persons other than those who had completed the questionnaires. Therefore, a different and broader perspective was obtained during the onsite interviews.

Since SAC did not choose to participate in the questionnaire phase of the survey but did participate in the onsite interviews, the level and character of data collected for this site differs from the other sites, although information was provided by participants during the visit to this site. This is particularly evident in the testing area, where sufficiently detailed data were lacking to make assessments about the general characteristics of the software testing environment at SAC. It should be noted that this is a more difficult site in which to characterize the testing environment because of the multimission nature of SAC responsibilities and the wide diversity of unrelated computing systems supported at this site.

2.3.1 Observations

The Air Force site summaries included in section 2.3.3 of this report provide objective data about those sites, to the extent that the survey participants were objective. In contrast to this information, certain subjective assessments about the nature of software environments at these sites were derived from the surveys. These assessments are included for the interest they may provide and are not necessarily substantiated by data.
a. The amount of the total project effort devoted to software development and maintenance requirements, implementation, testing, and installation was striking. Even in the hardware-oriented development programs, software development consumed a large share of the total effort. The impact of software processes on these systems is considerable and should not be underestimated.

b. There appeared to be relative uniformity among the Air Force sites in the application of software development methodologies. The sites surveyed generally employed system- and CPCI-level specifications, prepared them using MIL-STD-483 and -490, and used them for developing test requirements for system-level software testing. Also, software unit testing against design specifications was typically prepared according to MIL-STD-483 and -490. The sites made general use of design reviews (PDR and CDR) and adhered to Air Force regulations AFR 300-series and/or AFR 800-14. Recent technology for software development was used at the sites, including hierarchical design, program modularity, and incremental development techniques such as top-down or bottom-up integration. However, none of the sites were observed to be using the more recent, so-called advance techniques, such as formal specifications (InaJo, HDM), automated requirement languages (SREM/REVS), data-structured design techniques (Warnier, Orr, and Jackson), or formal testing techniques (symbolic execution). Of course, these omissions may be intentional, since these techniques do not lend themselves to application in these environments. Also, many of the surveyed computing systems have been in place for a number of years, preempting the application of such technology.

c. Both assembly language and compiler (HOL) debug aids are universally available; the compiler aids are widely applied and generally considered easy to use and effective. On the other hand, assembler debug aids seem to be somewhat avoided except where essential, probably due to their being more difficult to use.

d. Software testing aids are not widely available at the Air Force sites. The tools generally in use are directed at project test management and scheduling and at file and configuration management for testing purposes. Some specially developed tools were reported that aid in data manipulation and test data analysis, but they were infrequent. The most frequently used tools that directly support testing activities are environmental and system simulators. They also are used to support
development activities.

e. The prevalent methods of testing are debug (compile and remove coding errors), functional testing to specification requirements, and system operation, "soak." Methodical approaches toward establishing test goals and objectives for particular circumstances and shaping the test cases to those goals and objectives were not evidenced by the survey. There probably are no uniform practices for defining test cases among the Air Force sites.

f. There was a uniformly high level of concern among the Air Force sites about the adequacy of their software testing programs. All the sites directed considerable attention to testing; and the testing programs conducted either by the Government or by contractors are structured, progressive, and subject to management visibility. In all cases, final approval of testing is required before the software is permitted to become operational.

g. Documentation uniformly prepared for software testing programs consists of test plans, test procedures, and test reports. V&V plans are not used at all sites.

h. Independent V&V is not a common practice among the sites surveyed; it appears to be applied to large, complex contractor development programs. No instances of its use were noted where the software is developed or maintained by the Air Force sites surveyed.

i. Structuring of test programs at the Air Force sites shows a common symmetry. Testing consisted of specific, well-defined series of phases: debug and unit/module testing, often defined as computer program test and evaluation (CPT&E); integration and CPCI verification, often identified as DT&E; system testing and operational verification, often referred to as OT&E.

j. Analytical, metrical, statement, or logic coverage methods for determining the completion of testing are seldom used at the Air Force sites surveyed. The three most commonly used completion criteria are specification coverage, satisfaction of test requirements, and schedule completion.
2.3.2 Air Force Site Summaries

This section of the report provides summaries of the Air Force sites surveyed.

**Armament Division.** This site is a developing agency for tactical weapon systems, particularly threat, missile, and scoring systems. All embedded software systems development at the Armament Division (AD) is performed by contractors. The contractors are usually small, specialized, high-technology companies, but larger aerospace companies also contribute to the systems development. The software contained in these systems typically tends not to be critical, even though the systems themselves may be critical. The contractors design, develop, and test the software according to contractor-defined standards, but under the general contractual-level supervision of Air Force personnel. Testing practices vary rather widely among the many contractors supporting the AD.

**Aerospace Defense Center.** The Aerospace Defense Center (ADC) is both a developing and user agency for a single mission area: strategic warning and support systems. However, its systems are currently in place and undergoing maintenance activities, which consist of performance enhancements and additional system capabilities. Development is accomplished extensively by contractor personnel, depending on the organization and system function. Maintenance is either conducted by the Air Force or by contractors under close Air Force supervision. The system components are highly interrelated and the functions they perform tend to be highly critical. Systems are typically redundant, with multilevel fallback considerations. The computational systems perform numerous real-time communication processing functions, including message switching and routing control, complex trajectory calculations, systems status monitoring, and man-machine interface for control purposes. Testing practices are relatively uniform among the functional software activities and are highly adapted to the system characteristics. The space activity employs an independent test approach to interface and system-level testing after the programmer has completed module testing. Detailed test procedures are developed and updated prior to system testing. The requirement for successful software testing at both the module and system level is deeply embedded within the version release cycles. Operational testing to mission specifications is accomplished by the user subsequent to turnover from the software organization.
Aeronautical Systems Division. Aeronautical Systems Division (ASD) is a developing agency for weapon systems equipment, including avionics, automatic test equipment, crew training devices, flight control and reconnaissance, and C³ systems. System and software development are typically contracted. The development contractors tend to be medium-to large-size aerospace corporations, with substantial technical expertise in weapon systems development. The systems and embedded software are developed under well-defined contractual requirements and monitored by onsite representatives with frequent reviews of activities and documentation by ASD personnel. A wide diversity of software is developed by ASD, including numerous aircraft avionics and control systems, and communications systems software. Development activities are controlled by Government standards, and testing practices are fairly uniform, adhering to Air Force regulations and uniformly defined testing requirements imposed by SOW.

Strategic Air Command. The Strategic Air Command (SAC) has a diversity of missions to support (e.g., C³, war planning, intelligence support, and strategic weapons support) and develops a wide diversity of unrelated systems for these missions. For strategic weaponry, SAC is a user agency, while for the other areas it is both a developer and user. War planning and intelligence systems are developed and maintained almost exclusively by Air Force personnel, while the development of information and management systems often is conducted primarily by contractors and the maintenance shared by Air Force and contractor personnel. The software developed for the warning functions ranges from highly critical to noncritical. Software development practices for contractors are controlled by the SOW, and internal maintenance is conducted in accordance with SAC regulations. The computation systems used by SAC tend to be data-base and data-processing intensive, such as in the intelligence and war planning areas. The warning area includes real-time control functions, and the command centers use C³ technology software. SAC-conducted software testing practices and methods are standardized by SAC regulations. However, there exists variability in their application, corresponding to the differences in the software categories, criticality, and functional organizational practices.

Space Division. The Space Division (SD) is a development agency for space-related systems, including satellites, launch vehicles, and ground control and communications systems. SD relies extensively on contractors to develop its systems and the embedded software. These contractors also perform maintenance under follow-on contracts.
Software development requirements are defined in detail in the SOW; SD personnel, often coupled with technical consultant contractors, intensively monitor all development activities at all levels. Frequent reviews and technical direction are provided by this agency. A wide diversity of software categories is developed by SD, including software for communications, satellite control systems, prelaunch checkout and ground test systems, space vehicle avionics and control, and system simulations. This site employs independent verification and validation (IV&V) contractors to a greater extent than any of the other sites surveyed. Software testing practices are established by Air Force regulation and defined by the SOW. As a result, these practices tend to be relatively uniform among the development contractors. SD places great emphasis on the thoroughness, sufficiency, and formality of contractor testing practices.

**Tactical Air Command.** The Tactical Air Command (TAC) is the development and user agency for the major Air Force tactical planning system, the Computer-Assisted Air Force Management System (CAFMS). The CAFMS is a single-function, highly interrelated automated processing system. The major output product of CAFMS is the air tasking order report. CAFMS was developed by TAC personnel with some contractor assistance during the early requirement and design phases. Management, development, and maintenance of this system are well defined and uniquely adapted for its ongoing support. The system is currently operational, but undergoes continual enhancements and incorporation of new capabilities. The overall function of the CAFMS is quite critical, but few of its software components are considered to be more than moderately critical. The system does incorporate some automated fallback provisions in case of failure, but redundancy of system functions is not provided, and reversion to manual operation is the ultimate fallback provision. Testing practices are well defined and are incorporated as an integral part of a version release management system developed by TAC specifically for CAFMS. Testing is applied uniformly to all software components undergoing development.
3.0 SUMMARY OF TASK 2

3.1 EVALUATION OF TESTING TECHNIQUES

The relationship between mission application test requirements and state-of-the-art software test techniques was developed in this task. A step-by-step approach was applied in developing the tables used in the guidebook. This process began with developing the taxonomies to be used in developing the guidebook. First, the software error categorization was selected; second, the method was selected for organizing the types of software in the five USAF missions areas.

The first table constructed was the testing confidence level table. This table allows the guidebook user to establish a confidence level of software testing appropriate to the criticality of the software, the type of software, the budget and schedule constraints, etc. The testing confidence level number is then used in subsequent tables in the guidebook to extract appropriate testing tools and techniques for a specific situation.

The next table constructed rated the effectiveness of software test techniques for detecting specific software error types. This table is one of the few that does not use the testing confidence level. It uses an independent rating system that considers only the relative effectiveness of the specific tool or techniques being considered against specific software error types. This was followed by the construction of a table that related the tools and techniques to specific software types in terms of the testing confidence level. The software categories were derived from a survey of Air Force software testing practice. The relative criticality of the software, and its difficulty from a software engineering viewpoint were used to define the 18 categories used in the guidebook. These categories are listed in section 3.4 of this final report, and defined in table 2.2-2 of the guidebook.

A table that was to relate specific software error types to Air Force missions did not prove to be feasible. The task 1 survey of Air Force sites revealed that there was not enough data on software errors to construct the table. An attempt to provide a generic table was dropped when it became evident that such a table was too general; it contained no new or useful information.
A table that was to relate test phases to test techniques was reconsidered for its effectiveness. In some cases, the guidebook user may know his testing objectives, but in other cases the user may only know the relevant test phase. Therefore, this table was redesigned to relate test objectives (a more basic concern) to relevant test techniques. A new table was designed to relate test phases to test objectives. This change results in a more accessible and usable methodology for guidebook users, with varying kinds and levels of knowledge of their software environment and testing problems. A basic premise of this table, resulting from both state-of-the-art theory and discussions with Air Force test personnel, was that testing techniques used in early test phases remain applicable in later phases. For example, a code auditor, used in unit and module testing, should be used in later phases to assure that modifications and corrections resulting from later test phases do not violate coding conventions.

As a result of the discussion of the draft technical report on task 2 at an oral technical review, the last two tables were combined into one large table. The attendees agreed that the new combined table included in the guidebook was a more efficient and usable design than the two-table concept.

A table that relates software test techniques to support tools was constructed in a similar manner to the other tables. The major difference is that the table does not use the testing confidence level values as entries in the matrix. Instead, the support tools are merely indicated as being appropriate or inappropriate for the various test tools.

### 3.2 TASK 2 CONSIDERATIONS

Many considerations were involved in developing the contents of the tables. The ratings given in each entry reflect testing confidence level in terms of the complexity of the software, criticality of the software to the mission, and relevance of each software test tool or test technique to the particular software type.

The software categories were also chosen carefully so that the mission software types could be mapped into the software categories used in the guidebook. Within a category, software is chosen for similarity of function, internal structure, and complexity.
Two approaches were used to accomplish the complex task of evaluating the various testing techniques. First, a comprehensive literature search provided (1) a source of information on the various types of test techniques available and (2) a first cut at rating their effectiveness. Details on classes of testing techniques were gathered from articles on individual techniques, and relative ratings of technique effectiveness and limitations were gathered from survey articles. Second, a group of Boeing software engineers who had extensive experience in tools and testing were surveyed. This survey asked for the techniques to be rated according to each of the 10 considerations described in the following paragraphs.

a. Performance analysis capability. Determined by literature search, personal experience, and discussions with other software engineers.

b. Error detection and location capability. We rated each method, efficiency, relative success at detection of specific error types, and the precision with which the techniques locate (i.e., precisely identify the exact error) the software errors so that they can be understood, analyzed, and corrected efficiently.

c. Side effects and benefits. Other worthwhile results of the technique were also considered. For example, some techniques may provide output that can be used in the product documentation of the software being tested.

d. Cost and schedule benefits and impact. Specific techniques differ widely in their relative costs, the time involved in their use (schedule impact), and their relative benefit.

e. Management impact. All techniques were rated as to their benefit to management. Some approaches provide valuable visibility on the progress and health of the software development project. If a technique provided additional visibility to management, it was rated more valuable and productive (given a confidence level rating that would make its use more likely). It should be noted that some techniques such as a peer code review specifically exclude management visibility to achieve their goals.
f. Training required. Some techniques may be esoteric and require extensive training or personnel with special skills or backgrounds (e.g., algebraic and symbolic analysis). Other techniques involve common skills and are simple to apply (e.g., design and code walkthroughs).

g. User constraints. General rather than specific user constraints were considered in evaluating the techniques. The basic features characteristic of software test tools and techniques were used to determine the constraints imposed on the user. A tool may be available now only on the computers of a specific vendor, which limits the usefulness of that tool. However, in 6 months or a year, a similar tool may be available on many machines. In building rating tables for the guidebook, we considered the generic features, not the specific software test tools. The guidebook will provide considerations for evaluating specific tools and techniques.

h. Typical computer resources required. Each technique was evaluated in terms of its computer resources (time and storage) requirements compared to the potential benefits. Some techniques, such as peer code review, require little or no computer resources; whereas, other approaches (random testing or real-time testing) may use large blocks of time, large blocks of primary and secondary storage, or a considerable number of other computer resources.

i. Level of human interaction required. The various techniques considered in the tables vary considerably in the level of human interaction required. This human interaction must be considered at several levels: first, the amount of time the human must be involved compared to potential benefits; second, the degree of expertise required to effectively use the technique. The greater the proportional amount of time required or the level of expertise, the less attractive the technique was rated.

j. Usefulness of techniques in supporting modern programming practices. If the techniques encouraged the use of modern programming practices by the developers, or produced results useful for modern programming practices, they were rated more attractive.
3.3 SELECTION OF TOOL TAXONOMY

An understandable, comprehensive, user-friendly taxonomy of software test techniques was necessary for building the tables. Several taxonomies were evaluated as to their effectiveness according to the following criteria: they should be easy to understand, rational and systematic, free of inconsistencies, compatible with the intent of the guidebook, recognizable by the target audience, and compatible with the test phase relationship. The taxonomy has four major categories:

- Static analysis.
- Dynamic analysis.
- Symbolic testing.
- Formal analysis.

The first two categories include many software testing techniques. The complete taxonomy is described in detail in the following paragraphs.

Static Analysis. This is an automated analysis of computer program source code without executing the computer program. The major subcategories of static analysis are—

a. Code reviews and walkthroughs.
   1. Peer Review - review of code and design by project personnel.
   2. Formal Review - review by customer at scheduled points in the life cycle.

b. Error and anomaly detection techniques:
   2. Interface Checking - analysis of interface for consistency and completeness.
   3. Physical Units Checking - analysis of units of measure for consistency.
   4. Data Flow Analysis - analysis of sequences of program events to locate errors.

c. Structure analysis techniques and documentation:
   1. Structure Analysis - analysis of design or program structure to identify logical flaws.
   2. Documentation - production of documentation resulting from analysis (e.g., set-use listings).

23.
d. Program quality analysis:
   1. Halstead's Software Science - an attempt to formulate fundamental relationships for all computing languages.
   2. McCabe's Cyclomatic Number - an assessment of program complexity based on the number of branches.
   3. Software Quality Measures - a system to predict various software qualities (e.g., reliability, maintainability) based on a multitude of small, discrete measures, called metrics.

e. Input space partitioning. A path in a program consists of a possible flow of control. In path analysis, the input space is partitioned into path domains—those subsets of the program input domain that cause execution of the paths.
   1. Path analysis - detection of missing paths or incorrect paths.
   2. Domain testing - selection of test data on or near domain boundaries.
   3. Partition analysis - a method in which specifications of the program are partitioned into subspecifications that are then matched with domain partitions to generate a more sensitive test.

f. Data-flow guided testing. This is a method for obtaining structural information about programs (widely used for compiler design and optimization). One result is a set of dynamically meaningful relationships among program variables. Control flow information about the program is then used to construct test sets for the paths to be tested.

Dynamic Analysis. This is a method of analyzing a computer program in which the program itself is executed on a computer. The major subcategories of dynamic analysis are—
   a. Instrumentation-based testing. Programs are instrumented by statements or routines that do not affect the functional behavior of the program but record properties of the executing program.
   1. Path and structural analysis techniques - analysis of test coverage, execution frequency of branches, statements, etc.
2. Performance measurement techniques:
   (a) Timing and resource analysis - analysis of execution time and computer storage resources required by the software.
   (b) Algorithm complexity analysis - analysis of algorithm using a formalized approach.

3. Assertion checking - a technique using assertion statements in the source code that are then checked for validity during execution.

4. Debug aids - various facilities permitting trace, breakpoint, register inspection during or following execution.
   b. Random testing. This is a black-box technique in which a program is tested by randomly sampling inputs.
   c. Functional software testing. The specification of the program is viewed as an abstract description of its design and is then used as a guide to generate functional test data. Extremal and special values are the most important values in the input domain of a variable.
   d. Mutation testing. Mutation testing is a technique for evaluating test data adequacy. The program under test is changed (forming mutants of the original); test data are applied to the mutants. If the test data uncover the mutants, the data are accomplishing its job; if not, either the program is still correct in its mutated form, or the test data were inadequate to locate the mutant error.
   e. Real-time testing. This is the testing of software on "host" computers using environment simulators, as well as the testing of software on the "target" computer in the actual hardware or software system, or a simulation thereof.

Symbolic Testing. The input data and variables of a program are given formal or symbolic values, and the possible executions are characterized formally. The execution of the program is simulated by a symbolic evaluator that interprets the formal representation of the program and data.
Formal Analysis. This is a formal method of proving a design correct and uses rigorous mathematical techniques to analyze the algorithms of a solution. At present, formal analysis is primarily a manual activity with limited automated assistance.

This taxonomy provided the most useful working categorization for test techniques for the internal tables of the Software Test Guidebook.

3.4 SOFTWARE ENVIRONMENTS CHARACTERISTIC OF USAF MISSIONS

The task 1 survey provided information on the kinds of software characteristics of each USAF mission. A first draft of a table listing all the software types characteristic of each mission area was prepared, based on the task 1 interim technical report. This draft was submitted to the mission focal points for their review. The comments and criticisms were then used to build a final list.

There was much overlap between the mission areas, and a method of efficiently characterizing them for the guidebook was needed. Several software classification schemes were produced and evaluated. Guidelines used for evaluating classification schemes were the same as those used in selecting a tool taxonomy: the structure should be parallel in nature, the software types should be as unique as possible, and the terminology should be as clear as possible. A software classification scheme was chosen, and all the software types characteristic of the USAF missions were classified using that scheme. The classification of software categories follow.

- Batch (general).
- Event control.
- Process control.
- Procedure control.
- Navigation.
- Flight dynamics.
- Orbital dynamics.
- Message processing.
- Diagnostic software.
- Sensor and signal processing.
- Simulation.
- Database management.
- Data acquisition.
- Data presentation.
- Decision and planning aids.
- Pattern and image processing.
- Computer system software.
- Software development tools.
4.0 SUMMARY OF TASK 3

4.1 GUIDEBOOK DEVELOPMENT

Task 3 comprised two parts: first, design and prepare the guidebook; second, write a final report to describe the entire three-task research effort. The first part comprised about 95% of the task 3 effort and will be the only subject of this section.

The design of the guidebook was largely a matter of building a text structure to support the revised table structure from task 2. The guidebook preparation constituted two main efforts. The first effort was to prepare the text. One part of the text preparation was to write the directions and guidelines for tool selection. The second part of the text preparation was to assemble material that described all testing techniques and methodologies. The second effort was to prepare graphic material for the guidebook. This included the job of (1) rebuilding the guidebook tables to conform to the Air Force direction given at the task 2 oral report and (2) preparing graphic material to explain guidebook usage.

4.2 PREPARATION OF TEXT

The guidebook design was based on the table design completed in tasks 1 and 2. Several changes were made to the outline to structure the guidebook so that it is easy to use.

Directions and guidelines were written so that they are compact, clear, and easy to use. Thus, these sections are short and contain only general guidelines and considerations. Complete descriptions of techniques and their implementation were put into a separate section. This design frees the instruction sections of technique-specific information to make the methods of technique selection more obvious.

Most of the material for the section on technique descriptions was derived from two National Bureau of Standards publications: NBS Special Publications 500-93, "Software Validation, Verification, and Testing Techniques and Tool Reference Guide," (POW82A) and 500-98, "Planning for Software Validation, Verification, and Testing," (POW82B). This material was reviewed and updated as necessary.
4.3 PREPARATION OF GRAPHIC MATERIALS

As a result of the task 2 oral review, several major changes were made in the tables used for test technique selection. The basic categorization of testing techniques (taxonomy) was restructured. In addition, two of the tables in the original design were combined into one table.

The structured tables were compact and efficient to use, but difficult to understand at first glance. To correct this situation, a number of new diagrams were designed and built on the word processor. The diagrams explained in a step-by-step graphic manner how the guidebook and tables were to be used.

4.4 EDITING AND REVIEW

Technical editing was done first by the authors, who checked one another's material. The guidebook was then reviewed by the project manager. After completing the first draft, an independent technical editor was assigned to review the guidebook.
5.0 BIBLIOGRAPHY


(COR76) Cornall, L.M. and M.H. Halstead. "Predicting the Number of Bugs Expected In a Program Module", (Tech. Rep. CSD-TR-205). West Lafayette, IN: Purdue University, Computer Science Department, October 1976.


35.


(GOE81) Goel, Dr. Amrit. "A Guidebook for Software Reliability Assessment", under contract No. F30602-81-C-0169, for RADC.


(HEC81A) Hecht, H. "Synopsis of Interviews from a Survey of Software A Tool SoHaR Inc. report NBSIR 81-2388, November 1981.


37.


(KIN81) King, Bill. "ARGUS on BITS." Boeing Computer Services-SAMA Software Engineering Technology, November 1981.


(QUA83B) "Quality Metrics Framework Enhancements for Software Acquisition" (CDRL A003), RADC contract F30602-82-C-0137 with Boeing Aerospace Company, July 1983.


(STU81) Stucki, Leon G. "Using ARGUS on EKS II." Boeing Computer Services, March 1981.


(YOU77) Yourdon, E. Structured Walk-throughs. Yourdon, Inc. 1977.
MISSION
of
Rome Air Development Center

RARDC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C^3I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.